A Two-Factor Cointegrated Commodity Price Model with an Application to Spread Option Pricing*

Walter Farkas [°‖§] Elise Gourier [**] Robert Huitema [°] Ciprian Necula [°‡‡]

First version: October 15, 2015
This version: November 9, 2016

Abstract

In this paper, we propose an easy-to-use yet comprehensive model for a system of cointegrated commodity prices. While retaining the exponential affine structure of previous approaches, our model allows for an arbitrary number of cointegration relationships. We show that the cointegration component allows capturing well-known features of commodity prices, i.e., upward sloping (contango) and downward sloping (backwardation) term-structures, smaller volatilities for longer maturities and an upward sloping correlation term structure. The model is calibrated to futures price data of ten commodities. The results provide compelling evidence of cointegration in the data. Implications for the prices of futures and options written on common commodity spreads (e.g., spark spread and crack spread) are thoroughly investigated.

Keywords: Commodities, Cointegration, Futures, Option Pricing, Spread Options, Spark Spread, Crack Spread

JEL classification: C61, G11, G12

*We thank two anonymous referees as well as participants of the 2015 conference *Challenges in Derivatives Markets* held in Munich and the 2015 conference *European Financial Management Association (EFMA) Annual Meeting* held in Amsterdam for helpful comments and suggestions. We are also grateful to AXPO AG for sharing data.

°University of Zurich, Department of Banking and Finance, Plattenstrasse 14, 8032 Zurich, Switzerland
‖ETH Zurich, Department of Mathematics, Rämistrasse 101, 8092 Zurich, Switzerland and Swiss Finance Institute, Zurich, Switzerland
§Corresponding author. Email: walter.farkas@bf.uzh.ch
**Queen Mary University of London, Mile End Road, London E1 4NS, UK
‡‡Bucharest University of Economic Studies, Department of Money and Banking, Bucharest, Romania

One of the most distinctive features of commodity markets is the large number of long-run equilibrium relationships that exist between the levels of commodity prices. For example, the price of crude oil may move against the price of heating oil on a given day but not in the long run. That is, in the long-run the price of crude oil is tied to the price of heating oil in an equilibrium relation. These long-run equilibrium relations are usually referred to as *cointegration* relations. Cointegrated systems set in discrete time are widely employed in economics, especially empirical macroeconomics, to analyze various phenomena. Engle and Granger (1987) and Johansen (1991) revolutionized the field by a series of seminal results, such as the Granger representation theorem stating that a cointegrated system set in discrete time has an autoregressive error correction model (ECM) representation. Similar results are available for continuous time systems, but modeling cointegration in continuous time is less popular in economics. Phillips (1991) introduced the concept of cointegration in continuous time modeling and pointed out that the long-run parameters of a continuous time model can be estimated from discrete data by focusing on the corresponding discrete time ECM. Comte (1999) models cointegration in a continuous-time framework using CARMA (continuous-time autoregressive moving average) processes and develops a continuous time Granger representation theorem.

In this paper we develop a continuous time model of cointegrated commodity prices. In this model commodity prices are non-stationary and several cointegration relations are allowed amongst them.

There is a vast literature on modeling the price of a single commodity as a non-stationary process. For example, Schwartz and Smith (2000) assume the log price to be the sum of two latent factors: the long-term equilibrium level, modeled as a geometric Brownian motion, and a short-term deviation from the equilibrium, modeled as a zero mean Ornstein - Uhlenbeck (OU) process. To account for higher order autoregressive and moving average components in the short-run deviation from equilibrium, Paschke and Prokopczuk (2010) propose to model these deviations as a more general CARMA process. Moreover, Cortazar and Naranjo (2006) generalize the Schwartz and Smith (2000) model in a multi-factor framework. For an extensive account of the various types of one dimensional commodity price models, see the recent review of Back and Prokopczuk (2013). However, the literature on modeling a system of commodity prices is quite scarce.

Although cointegration has been studied extensively from a statistical and an econometric point of view, see, e.g., Baillie and Myers (1991), Crowder and Hamed (1993) and Brenner and Kroner (1995), it has received little attention in the continuous-time asset pricing literature.

Two fairly recent models, which are closly related to the one developed in this paper, are proposed in Cortazar et al. (2008) and Paschke and Prokopczuk (2009), both of which account for cointegration by incorporating common and commodity-specific factors into their modeling framework. Amongst the common factors, only one is assumed non-stationary. Although they explicitly take into account cointegration between prices, the cointegrated sytems generated by these two models are not covering the whole range of possible number of cointegration relations, but allow for none or for exactly $n-1$ relations to exist between the n prices. Benth and Koekebakker (2015) recently developed a model of two cointegrated commodites, where the deviations from the cointegration relation are modeled using a CARMA process.

Research on the pricing of derivatives written on cointegrated assets is understandably even scarcer. Duan and Pliska (2004) were (to the best of our knowledge) the first to examine the implications of cointegration for derivatives prices. They use a cointegrated multi-variate GARCH model to price options on the spread between two stocks. Their findings indicate that these prices only depend on the presence of cointegration when volatilities are stochastic. The recent paper by Nakajima and Ohashi (2012) also focuses on the implications of cointegration on spread options, but, in contrast to Duan and Pliska (2004), they consider commodity rather than equity spread options. Their modeling approach is also somewhat different from Duan and Pliska (2004). Nakajima and Ohashi (2012) write the continuous-time model of Gibson and Schwartz (1990) down in a multi-variate cointegration framework. More specifically, they assume that the risk-neutral drift of a commodity spot log-price temporarily deviates from the risk-free rate, and that these deviations are described by convenience yields and cointegrating relations among the prices. Their primary finding is that – unlike the result of Duan and Pliska (2004) – cointegration affects the prices of commodity spread options no matter whether volatility is stochastic or not. This discrepancy in the findings of Duan and Pliska (2004) and Nakajima and Ohashi (2012) is due to the fact that stocks (unlike commodities) earn the risk-free rate under the pricing (or risk-neutral) measure. This discrepancy is further discussed in Benth and Koekebakker (2015) and they propose a parametric class of pricing measures which preserves cointegration. Using their framework, a closed form formula for exchange options on two cointegrated commodities can be obtained, similar to the classical Margrabe's formula (Margrabe 1978).

Dempster et al. (2008) argue that when a cointegration relationship exists between two asset prices only the spread between the two should be modeled. This approach has the advantage that closed-form analytic pricing formulae for spread options may exist depending on the complexity of the dynamics. The authors obtain closed-form solutions under a two-

factor Ornstein-Uhlenbeck (OU) process for the dynamics of the spot spread. However, this "direct" approach also presents certain drawbacks: since only the spread is modeled, it is not possible to obtain prices for futures and options on the individual components of the spread, nor is it possible to use market data on these derivatives. Furthermore, it is not straightforward to formulate the dynamics of the spread because it may, for example, reach negative values, but certainly not large negative values. An OU or square-root process would then not be appropriate choices.

To avoid these disadvantages, we follow Duan and Pliska (2004) and Nakajima and Ohashi (2012), among others, and model the dynamics of the (underlying) assets forming the spread. More specifically, we extend the modeling framework developed by Schwartz and Smith (2000). They model the spot price dynamics of one commodity with two factors and show that their model is equivalent to the stochastic convenience-yield (SCY) model proposed by Gibson and Schwartz (1990). They argue, however, that their specification better conforms with intuition than models based on the somewhat elusive notion of "convenience yield". Our extension of the Schwartz and Smith (2000) model to n commodities is parsimonious and intuitive. Like Schwartz and Smith (2000), we use two factors to model the dynamics of the spot price of each commodity. Then, we assume that both factors are correlated across the n commodities, and that the long-term factor is also cointegrated across the n commodities.

The new feature of our model is that it can represent from one up to $n-1$ cointegration relationships, instead of only one (cf. Nakajima and Ohashi (2012)) or precisely $n-1$ (cf. Paschke and Prokopczuk (2009)).

The model is estimated using data of futures prices for crude oil, heating oil, gasoline, natural gas and electric power. Compelling evidence is found of multiple cointegration relationships. For example, our results indicate that two cointegration relationships exist between crude oil, heating oil, gasoline and natural gas. Furthermore, we find evidence for cointegration between natural gas and electric power, and between electric power in various markets. To gain a better understanding of our model and of its implications for futures and spread options, we calculate the prices of futures and options written on the *spark spread*, the *crack spread* and on a spread between electric power prices in different markets, using the results of our model estimation. The results give a clear and concise overview of the implications of cointegration in commodity markets. Confirming previous work (cf., e.g., Nakajima and Ohashi (2012)), we find that cointegration alters the shape of term-structures of futures prices at long horizons. Furthermore, we find that cointegration creates an upward-sloping correlation term-structure. The latter finding is consistent with empirical evidence,

but more importantly, it lowers the volatility of spreads, therefore lowering the prices of spread options.

The rest of the paper is organized as follows. In Section 1, we present the details of our model. Section 2 derives closed-form pricing formulae for futures and European-style options. In Section 3, we calibrate our model to futures price data on 10 energy commodities. The purpose of Section 4 is to demonstrate the effects of cointegration on futures prices and spread option prices. Section 5 concludes.

1 Commodity Spot Price Model

We consider n commodity spot prices $\boldsymbol{S}(t) = (S_1(t), \ldots, S_n(t))^\top$. The dynamics of each price is driven by two factors: $\boldsymbol{X}(t) = \log \boldsymbol{S}(t)$ captures the spot log-prices and $\boldsymbol{Y}(t) = (Y_1(t), \ldots, Y_n(t))^\top$ captures the long-run behaviour of $\boldsymbol{X}(t)$.[1] Intuitively, one can think of $\boldsymbol{X}$ and $\boldsymbol{Y}$ as the short-end and the long-end of the term-structure of futures prices.

Dynamics of log-prices

The seasonally adjusted spot log-prices follow the dynamics:

$$d(\boldsymbol{X}(t) - \phi(t)) = -K_x(\boldsymbol{X}(t) - \phi(t) - \boldsymbol{Y}(t))dt + \Sigma_x^{\frac{1}{2}} d\boldsymbol{W}_x(t), \tag{1}$$

where the $n \times n$ matrix K_x quantifies the speed of mean reversion of the adjusted log-prices around their equilibrium levels $\boldsymbol{Y}$, $\Sigma_x^{\frac{1}{2}}$ is a diffusion matrix and $\boldsymbol{W}_x$ is a standard Brownian motion. The function $\boldsymbol{\phi}(t)$ controls for seasonal effects and is defined as

$$\boldsymbol{\phi}(t) = \boldsymbol{\chi}_1 \cos(2\pi t) + \boldsymbol{\chi}_2 \sin(2\pi t), \tag{2}$$

where $\boldsymbol{\chi}_1$ and $\boldsymbol{\chi}_2$ are n-dimensional vectors of constants. Specifying the effects of seasonality as above is commonly done in the literature (see, e.g., Sørensen (2002), Paschke and Prokopczuk (2009) and Nakajima and Ohashi (2012))[2].

The long-run levels of log-prices have the following dynamics:

$$d\boldsymbol{Y}(t) = (\boldsymbol{\mu}_y - K_y \Theta \boldsymbol{Y}(t))dt + \Sigma_{xy}^{\frac{1}{2}} d\boldsymbol{W}_x(t) + \Sigma_y^{\frac{1}{2}} d\boldsymbol{W}_y(t), \tag{3}$$

[1] By modelling each commodity price by two factors, we have a total of $2n$ factors.

[2] A statistical study on the seasonal variability of energy commodity prices can be found in Manoliu and Tompaidis (2002), Borovkova and Geman (2006), Geman and Ohana (2009), among others. While not directly relevant for the pricing of options on (commodity) futures (as pointed out by Back et al. (2013)), a deterministic seasonal component may impact on model estimation (as pointed out by Lo and Wang (1995) and Back et al. (2013)), and hence indirectly on option prices.

where $\boldsymbol{\mu}_y \in \mathbb{R}^n$, K_y and Θ are $n \times n$ matrices, $\Sigma_{xy}^{\frac{1}{2}}$ and $\Sigma_x^{\frac{1}{2}}$ are diffusion matrices and $\boldsymbol{W}_y$ is a standard Brownian motion independent of $\boldsymbol{W}_x$.

Let us denote $\boldsymbol{Z}(t) = (\boldsymbol{X}(t) - \boldsymbol{\phi}(t), \boldsymbol{Y}(t))^\top$. The process $\boldsymbol{Z}$ is a multivariate Ornstein-Uhlenbeck (OU) process with dynamics:

$$d\boldsymbol{Z}(t) = [\boldsymbol{\mu} - K\boldsymbol{Z}(t)]\, dt + \Sigma^{\frac{1}{2}} d\boldsymbol{W}(t), \tag{4}$$

where we use the following vector and matrix notations:

$$\boldsymbol{\mu} := \begin{bmatrix} \mathbf{0}_n \\ \boldsymbol{\mu}_y \end{bmatrix}, \; K := \begin{bmatrix} K_x & -K_x \\ \mathbf{O}_n & K_y\Theta \end{bmatrix}, \; \Sigma^{\frac{1}{2}} := \begin{bmatrix} \Sigma_x^{\frac{1}{2}} & \mathbf{O}_n \\ \Sigma_{xy}^{\frac{1}{2}} & \Sigma_y^{\frac{1}{2}} \end{bmatrix}, \; \boldsymbol{W}(t) := \begin{bmatrix} \boldsymbol{W}_x(t) \\ \boldsymbol{W}_y(t) \end{bmatrix}. \tag{5}$$

The solution to (4) is given by:

$$\boldsymbol{Z}(T) = e^{-K(T-t)}\boldsymbol{Z}(t) + \left[\int_t^T e^{-K(T-u)} du\right] \boldsymbol{\mu} + \int_t^T e^{-K(T-u)} \Sigma^{\frac{1}{2}} d\boldsymbol{W}(u), \tag{6}$$

The conditional distribution of $\boldsymbol{Z}(T)|\boldsymbol{Z}(t)$ is a normal distribution with first two moments:

$$\begin{cases} E[\boldsymbol{Z}(T)|\boldsymbol{Z}(t) = \boldsymbol{z}_t] = e^{-K(T-t)}\boldsymbol{z}_t + \left[\int_t^T e^{-K(T-u)} du\right] \boldsymbol{\mu}, & (7) \\ Cov(\boldsymbol{Z}(T)|\boldsymbol{Z}(t)) = \int_t^T e^{-K(T-u)} \Sigma e^{-K^T(T-u)} du, & (8) \end{cases}$$

where $\Sigma = \Sigma^{\frac{1}{2}} \left(\Sigma^{\frac{1}{2}}\right)^T$ and the conditional variance-covariance matrix derives from Itô's isometry.

Cointegration

In the limit when $T - t$ goes to infinity, the conditional mean and covariance of $\boldsymbol{Z}$, given by (7) and (8), both go to infinity, unless all eigenvalues of the matrix K have positive real part. In that case, the distribution of the process stabilizes towards a normal distribution with unconditional first two moments given by:

$$\begin{cases} z_\infty = K^{-1}\boldsymbol{\mu}, & (9) \\ \text{vec}(\Sigma_\infty) = (K \oplus K)^{-1}\text{vec}(\Sigma), & (10) \end{cases}$$

where vec denotes the stack operator and $\oplus$ the Kronecker sum.[3] Therefore the two limiting

[3] The derivation of the covariance follows from Van der Werf (2007). Meucci (2009) studies the relationship between the multivariate Ornstein-Uhlenbeck process and statistical arbitrage and derives a similar formula.

first moments are finite and the process $\boldsymbol{Z}$ is stationary[4]. If some eigenvalues of K are zero, the corresponding eigenvalues of $e^{-K(T-t)}$ will be one and $\boldsymbol{Z}$ is no longer stationary. If $\boldsymbol{Z}$ is non-stationary, but there exist linear combinations of its components which are stationary, then it is called cointegrated. These linear combinations are referred to as cointegration relationships. From an economic perspective, cointegration exists when the paths of two or more variables (e.g., a pair or pool of stocks, interest rates or commodities) are linked to form a long-run equilibrium relationship from which they deviate only temporarily.

In our model, the matrix K summarizes the cointegration structure of the vector $\boldsymbol{Z}$. We can rewrite it as:

$$K = \begin{bmatrix} K_x & \mathbf{O}_n \\ \mathbf{O}_n & K_y \end{bmatrix} \begin{bmatrix} \boldsymbol{I}_n & -\boldsymbol{I}_n \\ \mathbf{O}_n & \Theta \end{bmatrix} = K_z \cdot \Theta_z,$$

where K_z can be seen as the speed of adjustment of $\boldsymbol{Z}$ towards its stochastic equilibrium level, defined through the cointegration matrix Θ_z as a linear combination of the components of $\boldsymbol{Z}$.[5]

Cointegration is included at two levels in our model. The first level of cointegration links the seasonally adjusted spot log-prices to their long-run levels. The corresponding cointegration relationship is described by the upper part of the Θ_z matrix: $[\boldsymbol{I}_n \quad -\boldsymbol{I}_n]$. We assume that K_x is invertible, and that all of its eigenvalues have positive real part. Hence there are n cointegration relationships between seasonally adjusted spot log-prices and their equilibrium levels. The matrix K_x quantifies the speed of mean reversion of the elements in $\boldsymbol{X}$ around the long term levels in $\boldsymbol{Y}$. The second level of cointegration links the components of $\boldsymbol{Y}$ to one and another through the lower part of the Θ_z matrix. We denote the number of cointegration relationships between them by h, where $h \geq 0$ and $h < n$. The respective $n \times n$ cointegration matrix Θ has the last $n - h$ rows equal to zero vectors and each of the first h non-zero rows encodes a stationary combination of the components of $\boldsymbol{Y}$. It is normalized such that $\Theta_{ii} = 1, i \leq h$. The matrix K_y is a $n \times n$ matrix with the last $n-h$ columns equal to zero vectors, such that $K_y\Theta$ is a $n \times n$ matrix of rank h, with all h non-zero eigenvalues having positive real part. Each of the h non-zero columns in K_y quantifies the speed of adjustment of each element in $\boldsymbol{Y}$ to the corresponding cointegration relation. We show in Appendix A that the vector process $\Theta_z \boldsymbol{Z}(t)$ is stationary, even if $\boldsymbol{Z}$ is not stationary.

In the empirical section, we will focus on the impact of the cointegration at the second

[4]Notice that since we will be working in a Gaussian framework, the notions of weak and strong stationarity are equivalent. Hence, for the (Brownian) Ornstein-Uhlenbeck process $\boldsymbol{Z}$, stationarity is guaranteed by finiteness of its first two limiting conditional moments.

[5]This decomposition, called full rank factorization, is not unique. Footnote 17 details how we have ensured identification.

level, captured by the Θ matrix. We can distinguish the following 3 cases:

(i) $\text{rank}(\Theta) = 0$ $\rightarrow$ no cointegration relationship exists among the (non-stationary) long-run log-price levels $\boldsymbol{Y}(t)$;

(ii) $0 < \text{rank}(\Theta) < n$ $\rightarrow$ cointegration among the long-run log-price levels $\boldsymbol{Y}(t)$ exists;

(iii) $\text{rank}(\Theta) = n$ $\rightarrow$ $\boldsymbol{Z}(t)$ is stationary.

In case (i) $\Theta = \mathbf{O}_n$ and if the diffusion matrices are diagonal we are effectively in the two-factor model of Schwartz and Smith (2000). Case (iii) corresponds to a so-called double-mean-reverting model.

To summarize, our model allows for h cointegration relationships between equilibrium log-price levels, with $0 \leq h < n$. By doing so, it extends, to our knowledge, the existing literature which usually restricts the number of cointegration relationships at the second level, to exactly one or $n - 1$. We will show in the empirical section that this additional level of flexibility is supported by the data.

To better illustrate the use of the model we briefly look at a toy example.

Example 1 (*cointegration matrix*)**.** Consider the following 3 possible configurations of the matrix Θ:

$$\text{(i)}: \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}, \qquad \text{(ii)}: \begin{bmatrix} 1 & -3 \\ 0 & 0 \end{bmatrix}, \qquad \text{(iii)}: \begin{bmatrix} 1 & -1 \\ 0 & 2 \end{bmatrix}.$$

Recall that each non-zero row of the matrix Θ corresponds to a cointegration relationship. This means that in case (i) there is no cointegration relationship. In case (ii) there is one cointegration relationship. The value of $Y_1(t)$ tends to stay close to 3 times the value of $Y_2(t)$. For case (iii), the matrix is invertible. In this case, $Y_1(t)$ and $Y_2(t)$ are both stationary. The value of $Y_1(t)$ tends to stay close to the value of $Y_2(t)$, which, in turn, tends to stay close to zero.

Closed-form solution

From equation (6), $\boldsymbol{X}(T)$ can be written as:

$$\begin{aligned} \boldsymbol{X}(T) = \boldsymbol{X}(t) + [\boldsymbol{\phi}(T) - \boldsymbol{\phi}(t)] + \begin{bmatrix} \boldsymbol{I}_n & \mathbf{O}_n \end{bmatrix} \left[e^{-K(T-t)} - \boldsymbol{I}_{2n} \right] \boldsymbol{Z}(t) \\ + \begin{bmatrix} \boldsymbol{I}_n & \mathbf{O}_n \end{bmatrix} \left[\int_t^T e^{-K(T-u)} du \right] \boldsymbol{\mu} + \begin{bmatrix} \boldsymbol{I}_n & \mathbf{O}_n \end{bmatrix} \int_t^T e^{-K(T-u)} \Sigma^{\frac{1}{2}} d\boldsymbol{W}(u). \end{aligned} \tag{11}$$

Then using the result by Carbonell et al. (2008) that

$$e^{-K\tau} = \begin{bmatrix} e^{-K_x\tau} & \psi(\tau) \\ \mathbf{O}_n & e^{-K_y\Theta\tau} \end{bmatrix} \quad \text{with} \quad \psi(\tau) := K_x \left[\int_0^\tau e^{-K_x(\tau-u)} e^{-K_y\Theta u} du \right], \tag{12}$$

we obtain after straightforward calculations the following result[6]:

$$\begin{aligned} \boldsymbol{S}(T) = \boldsymbol{S}(t) \exp \Bigg\{ & \left[e^{-K_x(T-t)} - \boldsymbol{I}_n \right] \boldsymbol{X}(t) + \psi(T-t)\boldsymbol{Y}(t) + \left[\phi(T) - e^{-K_x(T-t)}\phi(t) \right] \\ & + \left[\int_t^T \psi(T-u)du \right] \mu_y + \int_t^T \left[e^{-K_x(T-u)} \Sigma_x^{\frac{1}{2}} + \psi(T-u)\Sigma_{xy}^{\frac{1}{2}} \right] d\boldsymbol{W}_x(u) \\ & + \int_t^T \psi(T-u)\Sigma_y^{\frac{1}{2}} d\boldsymbol{W}_y(u) \Bigg\}. \end{aligned} \tag{13}$$

As we can see from (13), the vector of spot prices $\boldsymbol{S}(t)$ is an exponential function of Gaussian factors[7]. Hence the characteristic function of $\boldsymbol{X}(t)$ can be readily computed analytically, as we show in Lemma 1 in Appendix B.

Note that apart from the assumption of cointegrated spot log-prices, most of our model assumptions are fairly standard. This has several advantages. First, and most importantly, it allows us to cleanly isolate the effects of cointegration. Second, the futures prices are exponentially affine in the two factors and the prices of European-style options are still given by a Black-like formula (see Black (1976)), which enables fast and easy calculations. Also, we would like to stress that adding jumps by specifying a mean-reverting jump-diffusion process for $\boldsymbol{X}(t)$ would most likely not lead to different conclusions about the effects of cointegration. The reason is that, while cointegration inherently has a long-term effect, the effect of spot price jumps fades away in the long term. This is particularly the case of commodity spot price jumps, given their tendency to "mean-revert", see, e.g., Hambly et al. (2009).

Similarly to Schwartz and Smith (2000), our model can also be written as a convenience yield model. We provide the equivalent convenience yield model in Appendix C.

2 Futures and European-Style Option Prices

Futures and options are derivatives instruments with the spot as the underlying "stock". Our goal in this section is to compute the prices of these instruments (under the assumption of no-arbitrage). To that end recall the renowned risk-neutral valuation principle, which states that

[6] Note that the matrix exponential of A is denoted by e^A and the element-by-element exponential of A by $\exp(A)$

[7] Notice that our model belongs to the class of affine models proposed by Duffie and Kan (1996).

prices of derivatives are simply given by the expectation of their discounted terminal pay-off under a risk-neutral measure. In order to proceed, we need to perform a change of measure from the real-world measure to the risk-neutral measure. We assume that the spot commodity is not an investable asset[8], which causes the market to be "incomplete" irrespective of the stochastic process characterizing the spot returns. Therefore we need to specify the market prices of risk. We make the simplifying assumption that the market prices of risk are constant over time. This means that the change of measure transforms the Brownian motion $\boldsymbol{W}(t)$ to a Brownian motion $\boldsymbol{W}^*(t)$ plus a vector of constants, i.e.,

$$d\begin{bmatrix}\boldsymbol{W}_x^*(t)\\ \boldsymbol{W}_y^*(t)\end{bmatrix} = d\begin{bmatrix}\boldsymbol{W}_x(t)\\ \boldsymbol{W}_y(t)\end{bmatrix} + \begin{bmatrix}\boldsymbol{\lambda}_x\\ \boldsymbol{\lambda}_y\end{bmatrix} dt, \tag{14}$$

where $\boldsymbol{W}_x^*(t)$ and $\boldsymbol{W}_y^*(t)$ are standard Brownian motions under the risk-neutral measure (indicated by an asterisk), and where $\boldsymbol{\lambda}_x$ and $\boldsymbol{\lambda}_y$ are the market price of $\boldsymbol{W}_x(t)$ and $\boldsymbol{W}_y(t)$ risk, respectively.

The assumption of constant market prices of risk is in line with a large body of literature (see, e.g., Gibson and Schwartz (1990), Schwartz and Smith (2000), Lucia and Schwartz (2002), Kolos and Ronn (2008), Cortazar et al. (2008) and Benth and Koekebakker (2015)). While keeping the model parsimonious in the number of parameters, it implies that cointegration is preserved under the change of measure. Under such condition, the spot price is not a discounted risk-neutral martingale[9]. The validity of the assumption is assessed in Section 3.

We now readily obtain Proposition 1, which gives a closed-form formula for the futures prices of the n commodities.

Proposition 1. *At time t the vector of futures prices for contracts with maturity T is given by*

$$\boldsymbol{F}(t,T) = \exp\left\{\alpha(t,T) + \beta(T-t)\boldsymbol{X}(t) + \psi(T-t)\boldsymbol{Y}(t)\right\}, \tag{15}$$

[8]For example, in the gas and oil markets that we consider, one needs to be able to transport and store the commodity in order to speculate on it. The assumption that the spot commodity asset is not tradable is common in the literature, see, e.g., Gibson and Schwartz (1990), Schwartz and Smith (2000), Casassus and Collin-Dufresne (2005), Casassus et al. (2013) and Benth and Koekebakker (2015). It is generally believed to hold for most commodities, except for precious metals. If one, instead, assumes that the spot commodity is tradable, as Duan and Pliska (2004), then, the martingale condition implies that cointegration under assets disappears when changing from the real-world to the risk-neutral measure.

[9]Notice that a deviation from martingale behavior may be interpreted as a *convenience yield* for holding the spot commodity. See Appendix C for the equivalent convenience yield model (under $\mathbb{P}$).

with $\beta(\tau) := e^{-K_x\tau}$ *and with* $\alpha(t,T)$ *defined by*

$$\alpha(t,t+\tau) := \left[\phi(t+\tau) - e^{-K_x\tau}\phi(t)\right] - \left(\boldsymbol{I}_n - e^{-K_x\tau}\right) K_x^{-1}\boldsymbol{\mu}_x^* + \left(\int_0^\tau \psi(\tau-u)du\right)\boldsymbol{\mu}_y^* \quad (16)$$
$$+ \operatorname{diag}\left\{\frac{1}{2}\begin{bmatrix}\boldsymbol{I}_n & \boldsymbol{O}_n\end{bmatrix}\left[e^{-K\tau}\left(\int_0^\tau e^{Ku}\Sigma e^{K^\top u}du\right)e^{-K^\top\tau}\right]\begin{bmatrix}\boldsymbol{I}_n \\ \boldsymbol{O}_n\end{bmatrix}\right\},$$

where $\operatorname{diag}(A)$ *returns the vector with diagonal elements of* A*, and where*

$$\boldsymbol{\mu}^* = \begin{bmatrix}\boldsymbol{\mu}_x^* \\ \boldsymbol{\mu}_y^*\end{bmatrix} = \boldsymbol{\mu} - \Sigma^{\frac{1}{2}}\begin{bmatrix}\boldsymbol{\lambda}_x \\ \boldsymbol{\lambda}_y\end{bmatrix}.$$

Proof. See Appendix D. **q.e.d.**

From Proposition (1), we see that the futures price $\boldsymbol{F}(t,T)$ is an exponential affine function of the two factors $\boldsymbol{X}(t)$ and $\boldsymbol{Y}(t)$. Note that, while the coefficients β and ψ only depend on time t and maturity T through the (remaining) time to maturity $T-t$, α depends on both variables t and T independently (due to the seasonality function $\boldsymbol{\phi}(t)$).

Now by Itô's lemma the risk-neutral dynamics of $\boldsymbol{F}(t,T)$ are given by

$$\frac{d\boldsymbol{F}(t,T)}{\boldsymbol{F}(t,T)} = \left[e^{-K_x(T-t)}\Sigma_x^{\frac{1}{2}} + \psi(T-t)\Sigma_{xy}^{\frac{1}{2}}\right]d\boldsymbol{W}_x^*(t) + \psi(T-t)\Sigma_y^{\frac{1}{2}}d\boldsymbol{W}_y^*(t). \quad (17)$$

The variance-covariance matrix of returns on futures prices, denoted $\Xi(\tau)$, is an immediate consequence of (17). It depends on τ unless $K_x = \mathbf{O}_n$. Since (17) is a Gaussian-affine model, closed-form pricing formulae for European-style options can be readily obtained using Black's formula (see Black (1976)). It should be stressed that these pricing formulae have no explicit dependency on the market prices of risk (only indirectly via their dependency on $\boldsymbol{F}(t,\cdot)$). Indeed, the information on the market prices of risk is fully embedded in the term structure of futures prices. Hence, if options are priced using observable market-based futures prices, there is no reliance on the assumption about the market prices of risk.

3 Model Estimation

3.1 Data

The model is estimated on three sets of data. Dataset I consists of weekly time-series data on the closing futures prices of 4 closely linked energy commodities, namely crude oil, heating

oil, gasoline[10] and natural gas[11]. The contracts have fixed times to maturity of 1, 3, 6, 9 and 12 months and are listed and traded on the NYMEX (New York Mercantile Exchange). The data span the period from January 1991 to December 2013.[12]

Dataset II is composed of weekly time-series data on the closing futures prices of electric power in 4 different markets, namely Germany, Nordics,[13] Switzerland and Poland.[14] These data cover a relatively short time period from January 2007 to December 2013.

Typically the most liquid and high-volume electric power futures contracts are the so-called Base Month, Base Quarter and Base Year ones. The Base Month is a contract delivering 720 MWh (megawatt-hours) for a period of 1 month (1MW $\times$ 30 days $\times$ 24 h/day = 720MWh). Base Quarter (Year) contracts are nothing else than a series of 3 (12) consecutive Base Month contracts.

To make the model compatible with the data in set II, we regard 720 MWh electric power for a period of 1 month as a commodity, and we shall henceforth refer to it as the **power commodity**. Note that we have 4 power commodities, one for each region. The power commodity serves as the underlying asset for the Base Month contract. The theoretical (or model-implied) futures prices of the Base Month contract can be easily calculated by using (15). The Base Quarter contract with maturity T is a contract for which the seller commits to delivering power to the buyer, starting at T, for a period of a quarter. It is therefore, as noted earlier, a bundle of 3 Base Month contracts with maturities T, $T + 1M$ and $T + 2M$ (capital "M" means month). Similarly, the Base Year contract with maturity T is a bundle of 12 Base Month contracts with maturities T, $T + 1M$, ..., $T + 10M$ and $T + 11M$. The price of a Base Quarter (Base Year) contract can hence be approximated by the average[15] of the prices of 3 (12) Base Month contracts. For each type of contract (i.e., Base Month, Base Quarter and Base Year), we use in the measurement equation the futures contract with the nearest time-to-maturity. The reason is that, on some days and for some regions, only one maturity is available.

Dataset III contains weekly time-series data on the closing prices of electric power (Phelix

[10] From January 1991 to December 2005 data on Unleaded Gasoline (HU) is used, while data on Reformulated Blendstock for Oxygenate Blending (RB) is used for the remainder of the period (i.e., for the period January 1996 to December 2013). The reason is that the HU contracts stopped trading on NYMEX end of 2005. We note that the HU and RB contracts were trading at almost the same prices over the 2-year period they co-existed.

[11] This dataset was retrieved from Quandl (www.quandl.com).

[12] Each period, the futures prices with exactly 1, 3, 6, 9 and 12 months to maturity are interpolated from the existing term structure of futures prices.

[13] The Nordics consist of Sweden, Norway, Denmark, Finland, Iceland, The Faroe Island and Greenland.

[14] This dataset was obtained from the trading division of AXPO AG.

[15] Note that all prices are quoted in the same unit (€/MWh), independent of the length of the delivery period.

Base Futures) and natural gas (TTF Natural Gas Futures) futures contracts, both of which are traded on the EEX (European Energy Exchange)[14]. The data cover the time period from January 2006 to December 2013. The reason for not taking the weekly natural gas futures prices from dataset I and/or the weekly electric power futures prices from dataset II is simply because this data cannot be (straightforwardly) used to compute values of the *spark spread* (which is defined as the spread between the market price of electric power and the cost of the natural gas needed to produce that electric power). The "power commodity" in dataset III corresponds to electric power for one day and the "natural gas commodity" corresponds precisely to the amount of natural gas needed to generate that much electricity, which enables an easy calculation of the spark spread. Like for dataset II some futures contracts are conveniently represented as a bundle of 'basic' futures contracts with various expiration dates (with the "power commodity" or "natural gas commodity" as underlying) whose prices can be easily computed using equation (15).

3.2 Procedure

Since the state variables $\boldsymbol{X}(t)$ and $\boldsymbol{Y}(t)$ cannot be observed directly from the data, they are inferred from the (observable) futures prices, as previously employed in the literature (e.g. Cortazar et al. (2008), Paschke and Prokopczuk (2009), and Nakajima and Ohashi (2012)). We use the Kalman filter for this purpose. The state (or transition) equation is given by (1)-(3). The measurement equation is obtained by matching model-implied log-prices of futures to the data, accounting for noise by assuming that the error is an n-dimensional (uncorrelated) Gaussian process. We stress that the Kalman filter is applicable because the transition and measurement equations are subject to Gaussian noise and linear in $\boldsymbol{X}(t)$ and $\boldsymbol{Y}(t)$. The Kalman filter is also used for computing the likelihood of the (observed) futures prices given the model and the model parameters. This facilitates a straightforward maximum likelihood (ML) estimation of the model parameters.

Given that there are a large number of parameters, we employ a two-step ML estimation procedure. In the first step, we estimate a simple one-dimensional version of the model for each commodity separately. The estimates for the parameters of the seasonality term (2), as well as those for the variance of the measurement errors, are kept fixed in the second step. The estimates for the other parameters serve as starting point for the second step of the estimation procedure. The filtered values of the latent state variables $\boldsymbol{X}(t)$ and $\boldsymbol{Y}(t)$ are utilized to calculate reasonable starting values for parameters that cannot be estimated within the model version used in step one (i.e., correlation and cointegration parameters).

In the second step of the estimation procedure, the parameters of the full model are estimated (with exception of those that are kept fixed at values obtained in the first step). At this step, the log-likelihood function is maximized under the restriction that the model is "well-specified". Specifically, this means that at each iteration (of our maximization routine) we perform a Johansen constraint test,[16] checking whether the cointegration properties of the filtered time-series of $\boldsymbol{X}(t)$ and $\boldsymbol{Y}(t)$ do not contradict with the cointegration matrix Θ.[17]

3.3 Results

In this section, we present and discuss our estimation results.

Appendix E.2 illustrates the quality of the fit of log-prices of futures on crude oil, and shows that the data is remarkably well fitted. The ability of the model to fit the data is confirmed by Root Mean Square Errors (RMSEs) which are very low, ranging from 9 bps (1-year futures on crude oils and heating oil) to 1.03% (3-month futures on natural gas). These results give us confidence that the model reproduces the features of the data well. It is also interesting to notice that RMSEs tend to be lower for longer maturity contracts, which may be explained by potential spikes in the prices, whose impact on derivatives' prices might be substantial on the short-term, but are likely to disappear on longer horizons. This observation is reassuring, as the focus of our paper is on the long-term effects of cointegration.

In Appendix E.3 we collect graphs that represent the estimated trajectories of the spot log-prices $\boldsymbol{X}(t)$ and the long-run log-price levels $\boldsymbol{Y}(t)$. For all 10 commodities, the spot log-prices $\boldsymbol{X}(t)$ fluctuate quite wildly around the long-run log-price levels $\boldsymbol{Y}(t)$, which themselves are also varying over time (but with a lower magnitude, as expected). Furthermore, we find a strong seasonality in the (futures) prices of heating oil, gasoline and the 4 power commodities with the exception of "power Poland". The seasonality of crude oil prices is found to be negligible.

For comparison, the 10 long-run log-price levels $\boldsymbol{Y}(t)$ are also shown in Figures 1 through 3. Figure 1 suggests that the long-run log-price levels of crude oil, heating oil and gasoline are driven by only one factor. This would mean that no less than 2 cointegration relationships exist among these 3 variables. The long-run log-price level of natural gas, however, seems to be

[16] The Johansen test (see Johansen (1991)) is a procedure for testing the presence of cointegration relationships between two or more time-series.

[17] In order to obtain a unique full rank decomposition of the matrix K, one has to impose some extra structure on Θ and K_y by fixing some of their components. In our case, in addition to restricting the diagonal components of Θ to one, we also fix some components to zero. To achieve this goal, we employed an iterated estimation procedure that fixes to zero some of the parameters in Θ and K_y that are statistically non-significant in the previous step. Conditional on the structure of Θ and K_y, the decomposition is unique.

driven by another factor. These visual findings are confirmed by the estimated cointegration matrix (see (18)), denoted by $\hat{\Theta}_I$ for dataset I, by $\hat{\Theta}_{II}$ for dataset II, and by $\hat{\Theta}_{III}$ for dataset III.[18]

$$\hat{\Theta}_I = \begin{bmatrix} 1.00 & -0.69^{***} & -0.28^{***} & 0.03^{***} \\ 0.36^{***} & 1.00 & -1.47^{***} & 0.00 \\ 0.00 & 0.00 & 0.00 & 0.00 \\ 0.00 & 0.00 & 0.00 & 0.00 \end{bmatrix}, \quad \hat{\Theta}_{II} = \begin{bmatrix} 1.00 & -1.33^{***} & 0.00 & 0.00 \\ 0.00 & 1.00 & -0.63^{***} & 0.00 \\ 0.00 & 0.00 & 1.00 & -0.74^{***} \\ 0.00 & 0.00 & 0.00 & 0.00 \end{bmatrix}, \tag{18}$$

$$\hat{\Theta}_{III} = \begin{bmatrix} 1.00 & -2.27^{***} \\ 0.00 & 0.00 \end{bmatrix}.$$

It is clear from Figure 2 that the long-run log-price levels of power in the 4 regions are driven by only one factor. Intuitively this also makes sense because one would expect that price differences between regions (or countries) stabilize in the long-run. This visual finding is also confirmed by the estimated cointegration matrix $\hat{\Theta}_{II}$.

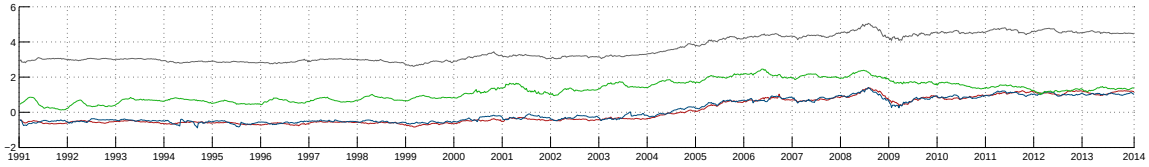

Figure 1. Filtered time-series of the long-run log-price levels $\boldsymbol{Y}(t)$ (dataset I). The colors corresponds to the following 4 commodities: gray (crude oil), red (heating oil), blue (gasoline), and green (natural gas).

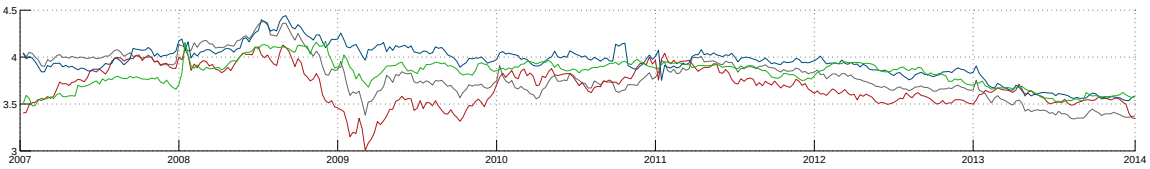

Figure 2. Filtered time-series of the long-run log-price levels $\boldsymbol{Y}(t)$ (dataset II). The colors corresponds to the following 4 commodities: gray (power Germany), red (power Nordics), blue (power Swiss), and green (power Poland).

For the sake of completeness, we also report the estimates of the parameters other than Θ. Values of these estimates can be found in Appendix E. These values are fairly standard and in line with the literature. Hence we do not explain them in detail but only mention a

[18] Statistical significance at 1% symbolized by ***, statistical significance at 5% symbolized by ** and statistical significance at 10% symbolized by *

Figure 3. Filtered time-series of the long-run log-price levels $\boldsymbol{Y}(t)$ (dataset III). The colors corresponds to the following 2 commodities: gray (electric power) and red (natural gas).

few results that we find interesting.

Recall that the matrices K_x and K_y measure the speed at which $\boldsymbol{X}(t)$ and $\boldsymbol{Y}(t)$ return to their cointegrated equilibria. From the estimates of K_x and K_y, we see that $\boldsymbol{X}(t)$ returns at least twice as fast to its equilibrium state than $\boldsymbol{Y}(t)$. This is a very intuitive result, because a deviation of $\boldsymbol{X}(t)$ from its "equilibrium" value is likely due to daily fluctuations in supply and demand (or speculative activity) whereas a deviation of $\boldsymbol{Y}(t)$ from its "equilibrium" value is more likely due to structural shifts in supply and demand.

A second observation is that $\boldsymbol{X}(t)$ and $\boldsymbol{Y}(t)$ are strongly positively correlated, as indicated by the estimates of Σ_{xy}. This result is again very intuitive if we recall that $\boldsymbol{X}(t)$ and $\boldsymbol{Y}(t)$ correspond to the short-end and the long-end of the term-structure of futures prices, which typically have a strong tendency to co-move (as evidenced by many empirical studies).

4 Implications for European-Style Spread Option Prices

In this section we analyze the prices of European-style options written on the price difference, or *spread*, between two or more commodities. Spread options present an ideal setting for investigating the implications of cointegration, since they crucially depend on the short- and long-run relation between the assets that make up the spread. Also, since spread options have become regularly and widely-used instruments in financial markets, for example for hedging purposes or for exploiting (statistical) arbitrage opportunities, there is a growing need for a better understanding of the effects of cointegration on their prices.

We will distinguish between options on spreads between 2 commodities and options on spreads between $n > 2$ commodities, such as the price difference between crude oil and a basket of various refined products. The reason is that for the 2-commodity case we have analytic (approximation) formulas, such as Margrabe's formula (see Margrabe (1978)) and Kirk's (approximation) formula (see Kirk (1995)), while for the general case (i.e. when $n > 2$) we have to rely on numerical methods[19].

[19] Also for the $n = 3$ case there are some (semi-)analytic approximation formulas, such as the extension of Kirk's formula proposed by Alos et al. (2011), but here we give preference to the Monte-Carlo technique.

4.1 The 2-Commodity Case

The holder of a European-style call option on the spread between 2 commodities receives at maturity T the pay-off

$$\max(S_1(T) - \alpha S_2(T) - k, 0), \tag{19}$$

where k is the strike price, and where $\alpha > 0$ is usually the ratio of the units of measure of $S_1(t)$ to the units of measure of $S_2(t)$ or, in case of an "input-output" (or production) spread (e.g., crack, dark or spark spread), the input-output conversion rate. When the strike price k equals zero a spread option is equivalent to an option to exchange one asset for another. In this special case, the price of the option can be computed analytically by Margrabe's formula (see Margrabe (1978)). In the general case, however, there are no closed-form pricing formulae for spread options (at least to the best of our knowledge). Instead, we need to rely on approximation formulae or extensive numerical computations.

For the 2-commodity case considered here we use the approximation formula of Kirk (1995) to calculate the price of a spread option with payoff (19). Note that Kirk's formula coincides with Margrabe's formula when the strike price k equals zero.

Alternative approximation formulae can be found in Carmona and Durrleman (2003) and Bjerksund and Stensland (2011). These formulae typically improve upon Kirk's formula in terms of numerical accuracy and efficiency, but they come at the cost of higher computational complexity. Since, for our purpose, the precision of Kirk's formula is sufficient, we keep the formulae simple and do not use any of these alternatives.

Let us start by looking at a toy example to grasp the impact of the cointegration parameters on the prices of futures contracts and the distributional properties of spreads.

Example 2 (*spreads between two commodities*)**.** Consider two commodities having a single cointegration relation between their long term components given by $Y_1 - \theta Y_2 = 0$ with $\theta > 0$. Hence the cointegration matrix is given by

$$\Theta = \begin{bmatrix} 1 & -\theta \\ 0 & 0 \end{bmatrix}.$$

Furthermore, assume that $\boldsymbol{\chi}_1 = \boldsymbol{\chi}_2 = \mathbf{0}_2$, $\boldsymbol{\mu}_y = \mathbf{0}_2$, $\boldsymbol{\lambda}_x = \boldsymbol{\lambda}_y = \mathbf{0}_2$ and

$$K_x = \begin{bmatrix} k_1 & 0 \\ 0 & k_2 \end{bmatrix}, \; K_y = \begin{bmatrix} l_1 & 0 \\ -l_2 & 0 \end{bmatrix}, \; \Sigma_x = \begin{bmatrix} \sigma_x^2 & 0 \\ 0 & \sigma_x^2 \end{bmatrix}, \; \Sigma_y = \begin{bmatrix} \sigma_y^2 & 0 \\ 0 & \sigma_y^2 \end{bmatrix}, \; \Sigma_{xy} = \mathbf{O}_2,$$

where k_1, k_2, l_1, l_2, σ_x and σ_y are positive constants.

At time t, the long-run distribution of $\boldsymbol{X}$, i.e., the distribution of $\boldsymbol{X}(t+\tau)$ for large τ, is Gaussian with mean $m_X(\tau) \approx \Psi_1 \boldsymbol{Y}(t)$ and variance $v_X(\tau) \approx \sigma_y^2 \Psi_2 \tau$. The matrices Ψ_1 and Ψ_2 are given by

$$\Psi_1 := \begin{bmatrix} \frac{l_2\,\theta}{l_1+l_2\,\theta} & \frac{l_1\,\theta}{l_1+l_2\,\theta} \\ \frac{l_2}{l_1+l_2\,\theta} & \frac{l_1}{l_1+l_2\,\theta} \end{bmatrix}, \quad \Psi_2 := \begin{bmatrix} \frac{\theta^2\left(l_1^2+l_2^2\right)}{(l_1+l_2\,\theta)^2} & \frac{\theta\left(l_1^2+l_2^2\right)}{(l_1+l_2\,\theta)^2} \\ \frac{\theta\left(l_1^2+l_2^2\right)}{(l_1+l_2\,\theta)^2} & \frac{l_1^2+l_2^2}{(l_1+l_2\,\theta)^2} \end{bmatrix}.$$

While the covariance matrices are diagonal, the presence of cointegration creates a positive correlation between the long-run behavior of the spot log-prices. One further sees that $X_1 - \theta X_2$ is stationary given that its mean and variance does not depend on τ. However, $X_1 - \alpha X_2$ with $\alpha \neq \theta$ is not stationary; its variance at time $t+\tau$ is, for large τ, approximately given by

$$\sigma_y^2 \left(\theta - \alpha\right)^2 \frac{1+l^2}{\left(l+\theta\right)^2} \tau$$

where $l := \frac{l_1}{l_2}$. The minimum variance variance is obtained when $l = \frac{1}{\theta}$.

While for the purpose of spread option pricing we are interested in the distributional properties of the spread between prices (rather than those of the spread between log-prices), it is worth noting that the mean and variance of $S_1 - \alpha S_2$ ($\alpha > 0$) can be readily obtained using the properties of the multi-variate log-normal distribution.

Now we fix the parameters as $\alpha = 1$, $\theta = 1.25$, $k_1 = k_2 = 2$, $\sigma_x^2 = 0.2$, $\sigma_y^2 = 0.05$ and

$$\boldsymbol{X}(0) = \begin{bmatrix} 3.69 \\ 2.95 \end{bmatrix}, \quad \boldsymbol{Y}(0) = \begin{bmatrix} 3.60 \\ 2.90 \end{bmatrix},$$

and then we run a simulation to assess the impact of the parameters l_1 and l_2 on the level and volatility of the futures spread with 10 year to maturity, as well as on the correlation between the log-returns on the two 10-year futures contracts and on the at-the-money (ATM) European-style call option written on the 10-year futures spread. The results are shown in Figure 4.

We see that if there is cointegration, the correlation between the spread components becomes stronger (reducing the volatility of the spread), resulting in lower spread option prices. That is, the price of the ATM spread option is almost 35% lower if $l_1 = l_2 = 0.1$ than it is if $l_1 = l_2 = 0$ (i.e. no cointegration). Figure 5 shows a comparison between the distribution of the spread 10 years out if $l_1 = l_2 = 0.1$ and the it if $l_1 = l_2 = 0$ (no cointegration), clearly

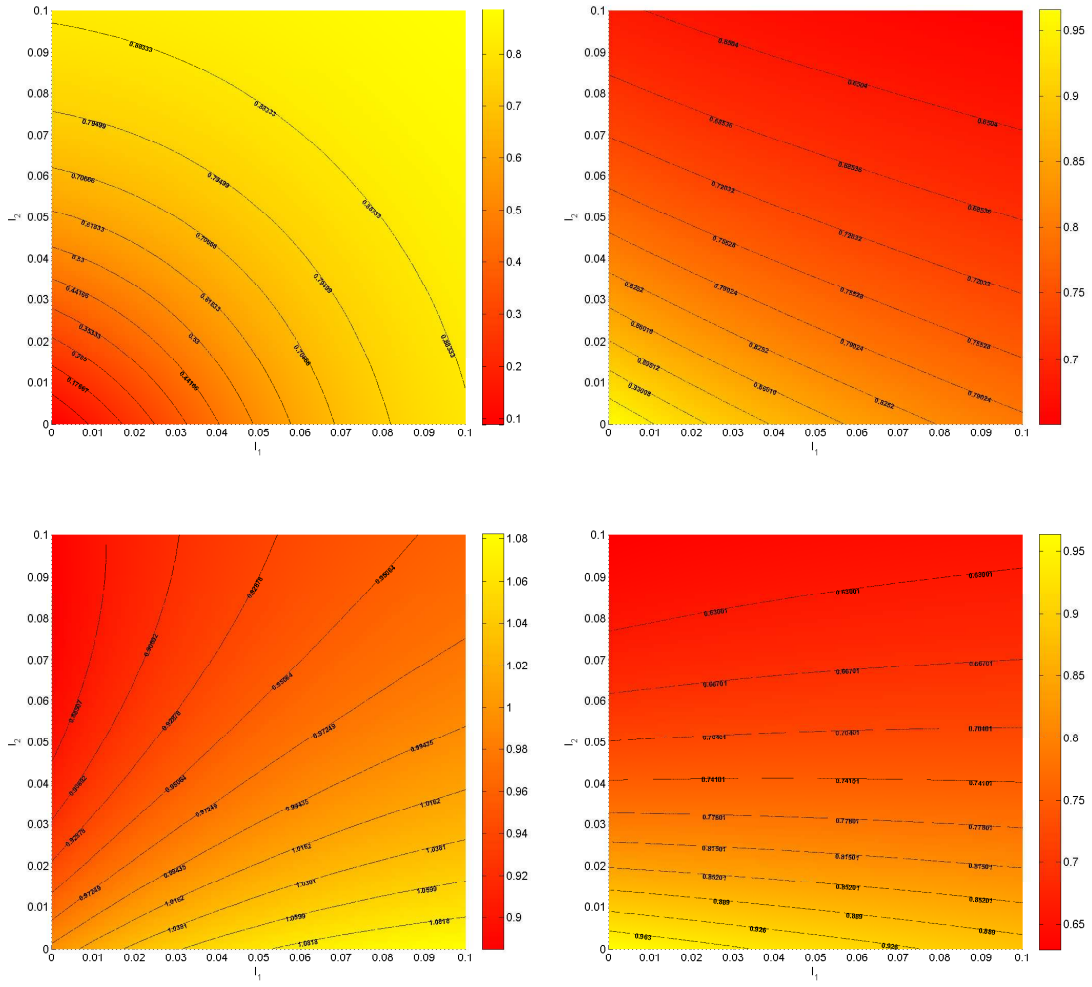

Figure 4. Top left panel: Correlation between log-returns on 10-year futures contracts. **Top right panel:** Kirk's (approximate) volatility of the 10-year spread. **Bottom left panel:** Level of the 10-year spread. **Bottom right panel:** Price of 10-year ATM call option on the spread. *Note:* Volatilities as well as futures and option prices have been rescaled to the case of no cointegration.

demonstrating that the width of the distribution decreases when cointegration is taken into account.

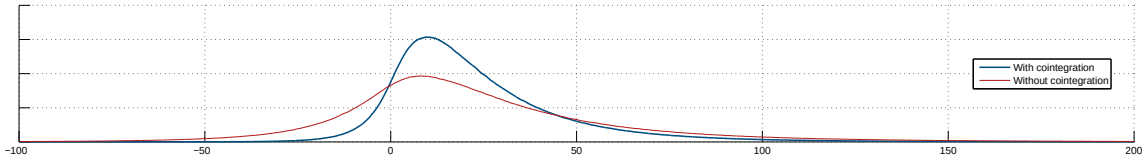

Figure 5. Distribution of the spread 10 years out.

As shown by Kirk's volatility in Figure 6, the impact of cointegration on the spread volatility is also visible on the shorter end of the term-structure. However, the effects

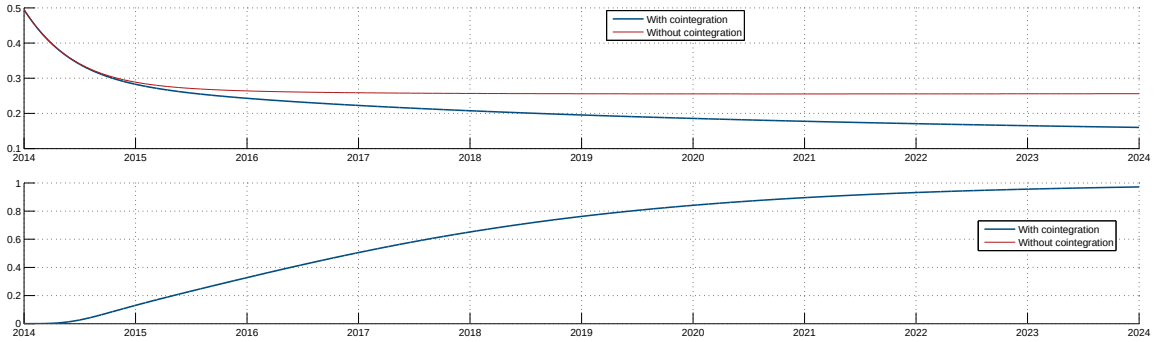

Figure 6. Top panel: Kirk's volatility $\widetilde{v}(t,T)$ as a function of maturity time T. **Bottom panel:** Correlation between the futures log-returns of both commodities. *Note:* The blue line corresponds to the original model, the red line to the model with no cointegration between the n variables in $\boldsymbol{Y}(t)$.

become negligible, and essentially disappear, for very short maturity spreads. The (significant) impact of cointegration on the volatility of long-dated spreads is (mainly) caused by stronger correlation between the log-returns on the futures contracts when their long-term price levels, i.e., $\boldsymbol{Y}(t)$, are cointegrated (see the bottom panel of Figure 6). This is an intuitive result given that in case of cointegration the far long-end of the spread term-structure is driven by only one factor. The term structure of futures prices and of ATM call option prices are shown in Figures 7 and 8. Figure 8 is the same as Figure 7, except that here no cointegration between the n variables in $\boldsymbol{Y}(t)$ is assumed.

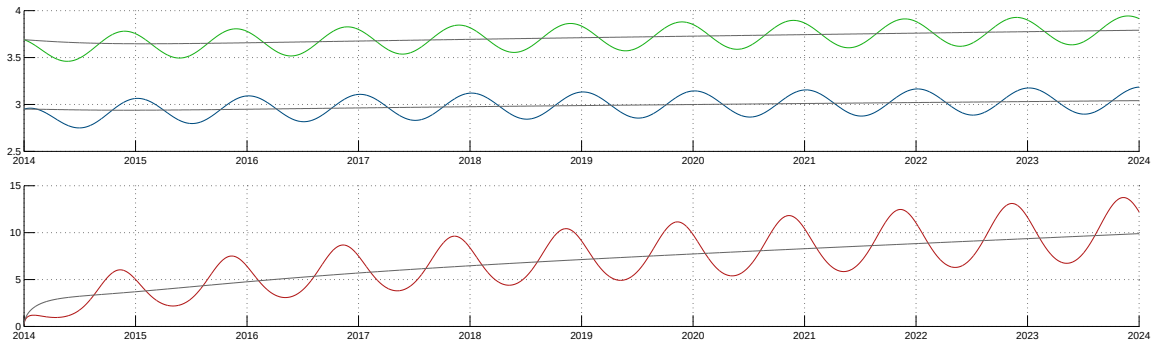

Figure 7. Top panel: Term-structure of futures log-prices for power (green line) and gas (blue line). **Bottom panel:** Term-structure of spark spread option prices. *Note:* The gray lines indicate the case without seasonality effects.

From Figure 7, we observe again a significant impact of seasonality on the results. This is the reason why we also plotted the results for the case without seasonality (depicted by gray lines). Let us now consider the shapes of the term-structures of futures prices. For that, it is important to note that these term-structures depict exactly how spot prices would evolve in absence of stochasticity (something that is also clear from the definition of futures prices, see (27)). As a consequence, the term-structures are subject to the same influences as the spot

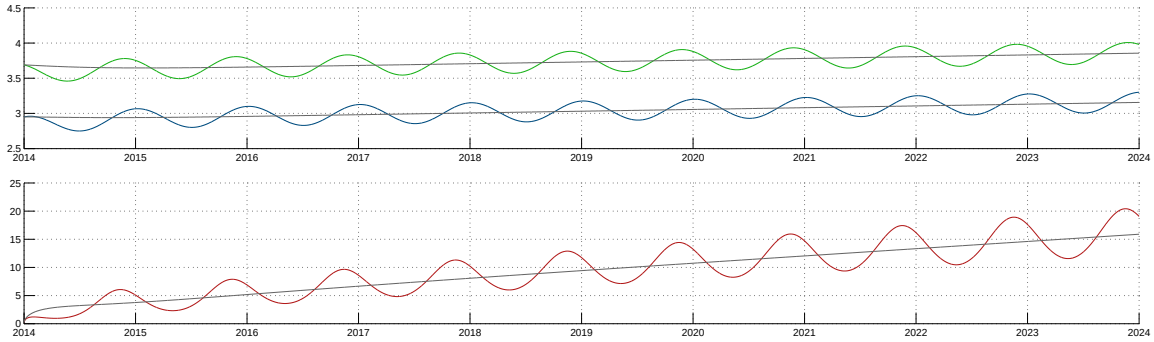

Figure 8. Top panel: Term-structure of futures log-prices for power (green line) and gas (blue line). **Bottom panel:** Term-structure of spark spread option prices. *Note:* The gray lines indicate the case without seasonality effects.

prices $\boldsymbol{S}(t)$, namely (i) the reversion of the seasonally adjusted spot log-prices $\boldsymbol{X}(t) - \boldsymbol{\phi}(t)$ towards their long-run levels $\boldsymbol{Y}(t)$ and (ii) the reversion of the latter towards their cointegrated "equilibrium" relationship(s) (characterized by the matrix Θ). Since the speed of reversion "(i)" is greater than that of reversion "(ii)" (see also the discussion in Section 3 above), "(i)" primarily influences the short-end of the term structures while "(ii)" primarily influences the long-end of the term structures. This is also clearly reflected in the figures: Up to a maturity of one year the shape of the term-structures is clearly determined by "(i)", that is, the one-year seasonally adjusted futures log-prices are relatively close to $\boldsymbol{X}(t = 2014) = (3.69, 2.95)^\top$. For maturities longer than one-year we see a convergence towards the cointegration "equilibrium" levels, i.e., $K_y \Theta \log \boldsymbol{F}(t = 2014, T = 2024) \approx \boldsymbol{\mu}_y$. This also explains why the term-structures in Figure 7 are very similar (to the ones in Figure 8) up to $T = 2015$ but differ for larger T. Spread option prices are plotted in the bottom panel of Figures 7 and 8. An immediate observation is that those in Figure 7 (with cointegration in $\boldsymbol{Y}(t)$) are substantially lower than those in Figure 8 (without cointegration in $\boldsymbol{Y}(t)$) and this is, as discussed above, mainly due to a lower variance of the spread because of positive correlation induced by cointegration.

We will now present another example on spreads between two commodities in which a so-called *spark spread* option is priced. In this example, we rely on the estimation results for dataset III (see Section 3).

Example 3 (*spark spread*)**.** The spark spread is, as mentioned, the spread between the market price of electric power and the cost of the natural gas needed to produce that electric power. As such, the spark spread is a metric of the profitably of natural gas-fired electric power plants. Hence, utility companies employ the spark spread as an indicator for turning on or off their natural gas-fired electric power plants. When this is not possible (in the short-term), they often resort to buying (financial) futures and/or options on the spark spread to hedge their

risk of a profit margin squeeze. In this example, the prices of these two types of instruments are calculated using the estimation results of dataset III.

The spot price of the spark spread at January 1st 2014 (according to our model) is \$8.87. Using Kirk's formula, we now compute the price of an at-the-money (ATM), hence $k = 8.87$, European-style call option with maturities up to 10 years starting January 1st 2014. For the sake of clarity we set the vector of risk premiums $\boldsymbol{\lambda}_x$ and $\boldsymbol{\lambda}_y$ and the risk-free interest rate curve equal to zero. The results are shown in Figures 9 and 10. Figure 10 is the same as Figure 9, except that here no cointegration between the n variables in $\boldsymbol{Y}(t)$ is assumed.

The shapes of the term-structures of spread futures prices and that of the ATM option prices are similar in the two cases since the estimated speeds of reverting to the cointegration relation (i.e. the elements of K_y) are both close to zero and not statistically significant. The effect of cointegration is most visible in the term-structure of the correlation of the futures log-returns between the two components of the spark spread, depicted in Figure 11.

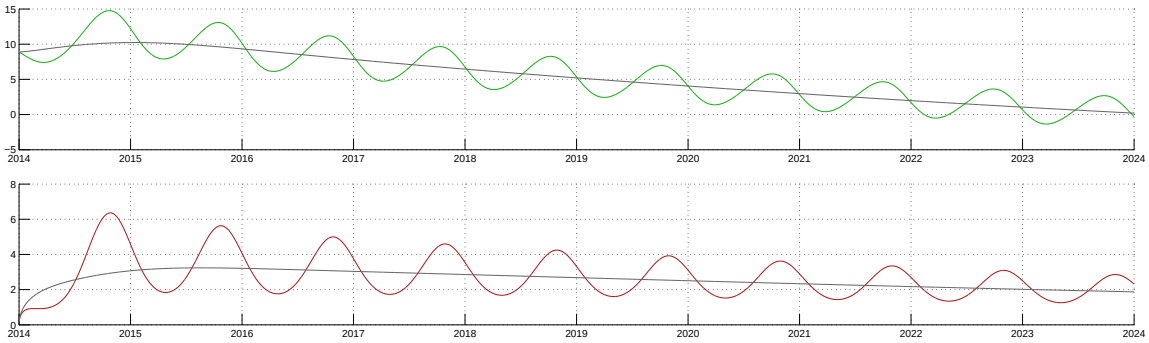

Figure 9. Top panel: Term-structure of spread futures. **Bottom panel:** Term-structure of spark spread option prices. *Note:* The gray lines indicate the case without seasonality effects.

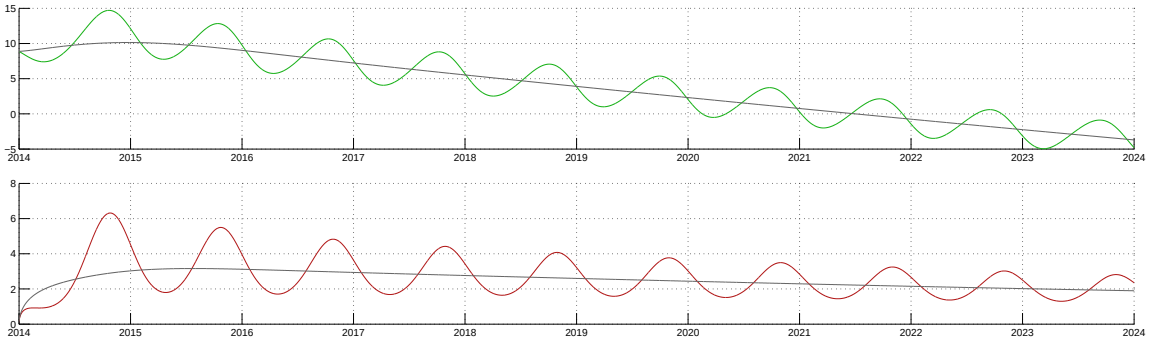

Figure 10. Top panel: Term-structure of spread futures. **Bottom panel:** Term-structure of spark spread option prices. *Note:* The gray lines indicate the case without seasonality effects.

In contrast to the previous example, the correlation of the futures log-returns is strictly positive even in the absence of cointegration. This is due to the fact that the estimated

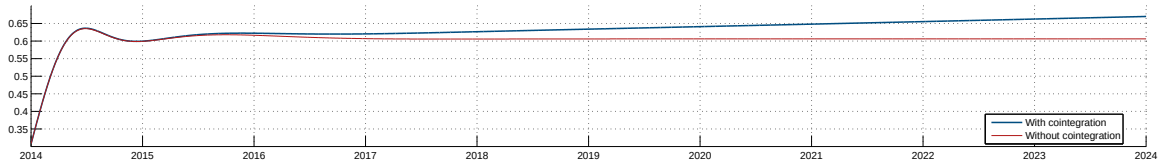

Figure 11. Correlation between the futures log-returns of both commodities. *Note:* The blue line corresponds to the original model, the red line to the model with no cointegration between the n variables in $\boldsymbol{Y}(t)$.

covariance matrix of the shocks is not diagonal. However, cointegration induces additional correlation that is increasing with maturity.

4.2 The n-Commodity Case

There is an extensive literature on approximation methods for spread and basket options on more that two commodities, with recent contributions from Li et al. (2010) and Caldana and Fusai (2013). However, mostly for simplicity, we rely in this paper on the Monte-Carlo simulation method for pricing spread options written on more than two commodities.

From (17), it follows that $\boldsymbol{F}(t,T)$ (conditional on information available up to time $s \leq t \leq T$) is distributed as follows:

$$\boldsymbol{F}(t,T) \sim \log \mathcal{N}\left(\log \boldsymbol{F}(s,T) - \frac{1}{2}\int_s^t \operatorname{diag}(\Xi(T-u))du, \int_s^t \Xi(T-u)du\right), \tag{20}$$

where $\operatorname{diag}(X)$ denotes the vector containing the diagonal elements of the matrix X. Note that $\boldsymbol{F}(s,T)$ can be either computed from (15) or observed from data.

The fact that the distribution function of $\boldsymbol{F}(t,T)$ is known in an easy-to-use and analytic form is one of the merits of our model. It allows us to simulate futures price curves at any time t in the future based on today's curves (time s) almost effortlessly. Hence, the price of a call option on the time-T value of a certain spread can be simply obtained by carrying out the following steps:

(i) compute or observe today's futures price curves $\boldsymbol{F}(s,T)$;

(ii) compute M realisations $\boldsymbol{F}^{(m)}$ $(m = 1,\ldots,M)$ of $\boldsymbol{F}(T,T)$ by sampling from (20) as follows:

$$\boldsymbol{F}^{(m)} = \boldsymbol{F}(s,T)\exp\left\{\boldsymbol{\varepsilon}^{(m)}\right\},$$

where $\boldsymbol{\varepsilon}^{(m)}$ is generated from a multivariate normal distribution with mean $-\frac{1}{2}\int_s^T \operatorname{diag}(\Xi(T-u))du$ and variance-covariance matrix $\int_s^T \Xi(T-u)du$;[20]

[20] Here the technique of antithetic variables is used to reduce the number of random samples needed for a

(iii) compute the Monte-Carlo estimate of a call with strike k on the spread

$$S_1(T) - \sum_{n=2}^{N} \omega_n S_n(T) \qquad \left(= \; F_1(T,T) - \sum_{n=2}^{N} \omega_n F_n(T,T) \right),$$

with $\omega_n \in \mathbb{R}_+$ for all $n = 2, \ldots, N$, as follows:

$$\frac{1}{M} \sum_{m=1}^{M} \max \left\{ F_1^{(m)} - \left[\sum_{n=2}^{N} \omega_n F_n^{(m)} \right] - k, 0 \right\}. \tag{21}$$

We note that the random variables $\boldsymbol{\varepsilon}^{(m)}$ can be simply re-used for pricing spread options with different maturity dates.

Example 4 considers the pricing of an option on the *crack spread*, which is a spread between the futures prices of 3 commodities. This example will rely on the estimation results for dataset I (see Section 3).

Example 4 (*crack spread option*)**.** The crack spread refers in general to the price difference between crude oil and refined products such as diesel, kerosine, gasoline and petroleum. The crack spread is hence nothing else than the profit margin that an oil refiner realizes when "cracking" crude oil while simultaneously selling the refined products in the wholesale market.

In this example we will consider a so-called 3:2:1 crack spread, which consists of shorting 3 crude oil futures contracts and purchasing 2 gasoline futures contracts and 1 heating oil futures contract. Since gasoline and heating oil are quoted in dollars-per-gallon while crude oil, and usually also the crack spread, is quoted in dollars-per-barrel, we must multiply the futures prices of gasoline and heating oil by 42 (the number of gallons per barrel). The crack spread is therefore calculated as follows:

$$-3 \times F_{\text{crude oil}}(t,T) + 2 \times 42 \times F_{\text{gasoline}}(t,T) + 42 \times F_{\text{heating oil}}(t,T). \tag{22}$$

An oil refiner can hedge the risk of losing profits by buying an appropriate number of futures contract on the crack spread. However, then he also loses any upside. As an alternative the oil refiner can buy a call option of the crack spread. Then he pays a fixed up-front cost (premium), but still profits from any future widening of the spread. In the following, we will calculate the prices of futures and call options on the crack spread using the estimation results of dataset I.

The spot price of the spread at January 1st 2014 (according to our model) is \$65.52. Now

given level of accuracy.

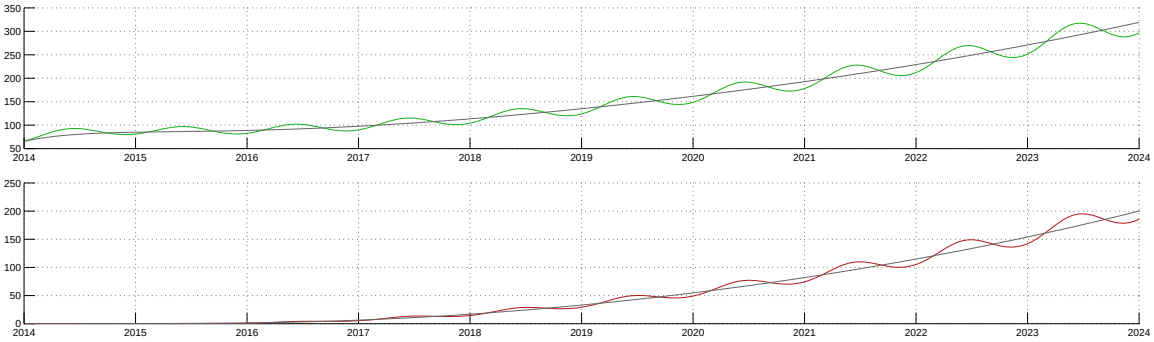

Figure 12. Top panel: Term-structure of futures crack spread prices. **Bottom panel:** Term-structure of crack spread option prices. *Note:* The gray lines indicate the case without seasonality effects.

we price an (out-of-the-money) European-style call option with strike $K = \$250.00$ on the crack spread with maturities up to 10 years starting January 1st 2014. For this purpose we rely on the Monte-Carlo based method presented above. The results are summarized in Figures 12 through 14. Like Figure 10, Figure 13 shows the case where there is no cointegration between the variables in $\boldsymbol{Y}(t)$.

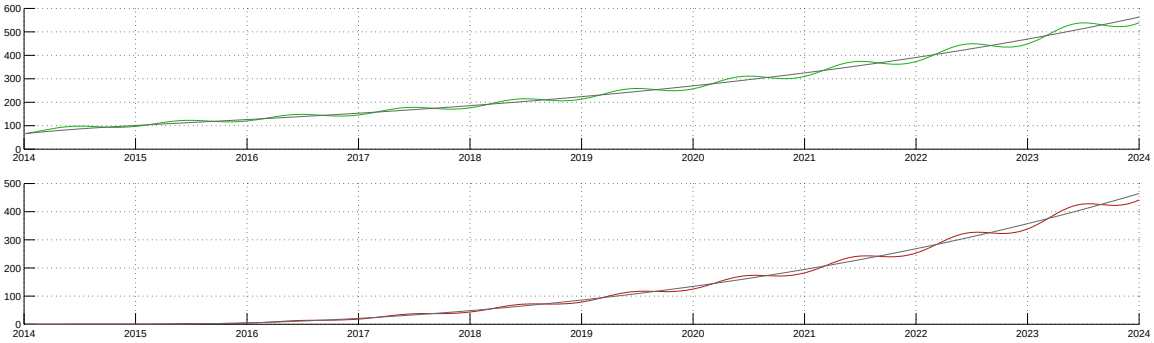

Figure 13. Top panel: Term-structure of futures crack spread prices. **Bottom panel:** Term-structure of crack spread option prices. *Note:* The gray lines indicate the case without seasonality effects.

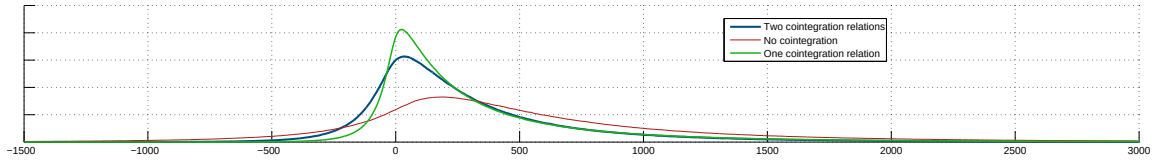

Figure 14. The distribution of the crack spread 10-years out.

The spread option prices are significantly lower when cointegration is accounted for. To better understand the latter observation, we have plotted in Figure 14 the probability distribution of the spread 10 years out. We find that the shape of the distribution is wider

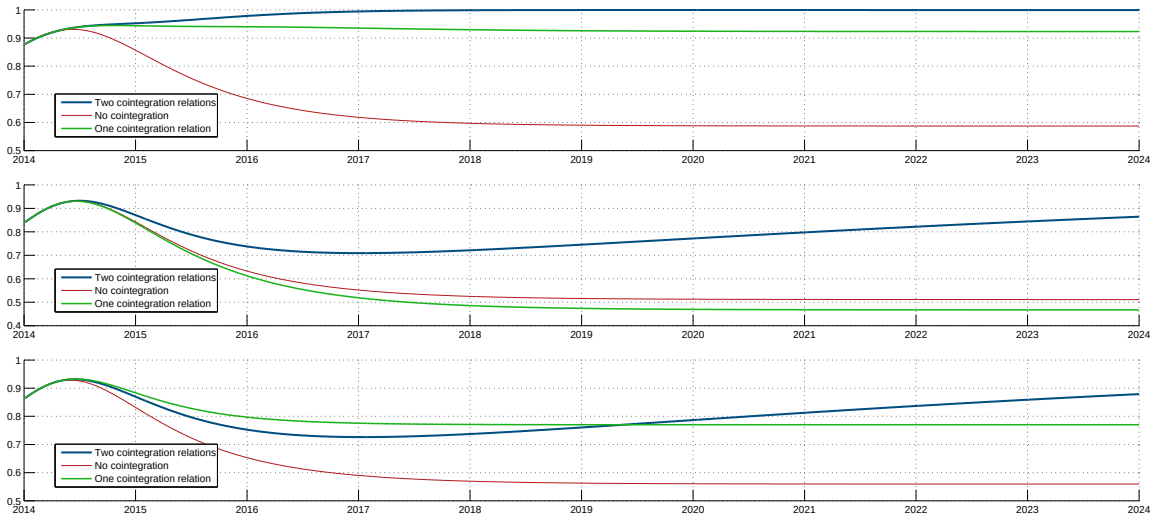

Figure 15. Correlation between the futures log-returns of three commodities (from top to bottom: between 1 and 2, between 2 and 3, between 1 and 3). *Note:* The blue line corresponds to the original model, the red line to the model with no cointegration between the n variables in $\boldsymbol{Y}(t)$, and the green line to the model with only the first cointegration relation between the n variables in $\boldsymbol{Y}(t)$.

without cointegration and gets narrower when the cointegration relations are included in the model. The variance of the crack spread distribution is lower in the model with cointegration given higher positive correlations between the three components. Indeed, Figure 15 depicts the effect of cointegration on the term-structure of the correlation of the futures log-returns between the three components of the crack spread. Again, the correlation of the futures log-returns is strictly positive even in the absence of cointegration. However, the correlations in the model without cointegration quickly reduces during approximatively the first three years. On the other hand, the additional correlation induced by cointegration is increasing with maturity.

Finally, we present an example of a spread between four commodities. In this example, we rely on the estimation results for dataset II (see Section 3).

Example 5 (*power spread*)**.** Consider a spread between the price of electric power on the Swiss market and the average between the prices of electric power in the German, Nordic and Polish markets:

$$F_{\text{Switzerland}}(t,T) - 0.33 \times \left(F_{\text{Germany}}(t,T) + F_{\text{Nordics}}(t,T) + F_{\text{Poland}}(t,T)\right). \tag{23}$$

The spot price of the spread at January 1st 2014 according to our model is 13.19.

Figure 16 depicts the distribution of the spread 10 years in the future. Again, the shape of the distribution is wider if cointegration is not taken into consideration. This has of course a

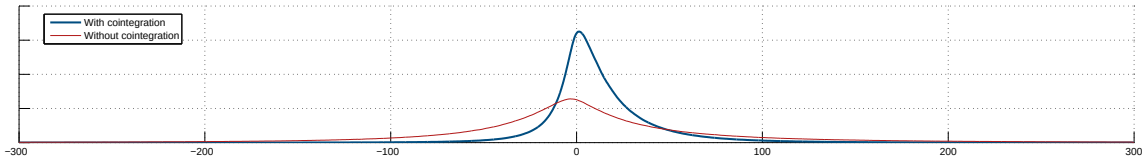

Figure 16. The distribution of the power spread 10-years in the future.

significant impact on option prices. For example, the price of a ATM call on the spread with 10 years to maturity is three times higher in the model without cointegration. As in the previous examples, the lower variance of the spread distribution in the model with cointegration is due to higher positive correlations between the four components.

5 Conclusion

We present a parsimonious two-factor model for n cointegrated commodity prices. The spot prices are fully cointegrated with the corresponding long-run levels, which are cointegrated with one another, across commodities. Expanding on Paschke and Prokopczuk (2009) and Cortazar et al. (2008), our model is able to describe one up to $n-1$ cointegration relationships between the long-term factors of n commodities.

We estimate our model to three alternative datasets and show that the data supports our model assumptions on cointegration. First, we use weekly prices of futures on crude oil, heating oil, gasoline and natural gas with a wide range of maturities, spanning a period of over twenty years. We find two (statistically significant) cointegration relationships. Only natural gas is not cointegrated with one of the others. Second, we use weekly prices of futures on electric power in four different regions. As expected, we find that all of them are strongly cointegrated with one another. The third dataset consists of futures prices on electric power and natural gas, which are also found to be cointegrated.

Using the estimation results of our model, we calculate the prices of several spread options on energy commodities and highlight the effects of cointegration. We show that cointegration affects particularly the long-end of futures price term-structures. Furthermore, we find that cointegration leads to an upward sloping correlation term-structure which lowers the volatility of spreads and therefore also lowers the value of options on spreads.

References

Alos, E., Eydeland, A., and Laurence, P. (2011). A Kirk and a Bachelier formula for three asset spread options. *Energy Risk*, 9(2011):52–57.

Back, J. and Prokopczuk, M. (2013). Commodity price dynamics and derivative valuation: A review. *International Journal of Theoretical and Applied Finance*, 16(6):1350032.

Back, J., Prokopczuk, M., and Rudolf, M. (2013). Seasonality and the valuation of commodity options. *Journal of Banking & Finance*, 37:273ñ290.

Baillie, R. T. and Myers, R. J. (1991). Bivariate garch estimation of the optimal commodity futures hedge. *Journal of Applied Econometrics*, 6(2):109–124.

Benth, F. E. and Koekebakker, S. (2015). Pricing of forwards and other derivatives in cointegrated commodity markets. *Energy Economics*, (52):104ñ117.

Bjerksund, P. and Stensland, G. (2011). Closed form spread option valuation. *Quantitative Finance*, (ahead-of-print):1–10.

Black, F. (1976). Studies of stock price volatility changes. In *Proceedings of the 1976 Meetings of the American Statistical Association*, pages 171–181.

Borovkova, S. and Geman, H. (2006). Seasonal and stochastic effects in commodity forward curves. *Review of Derivatives Research*, 9:167 – 186.

Brenner, R. J. and Kroner, K. F. (1995). Arbitrage, cointegration, and testing the unbiasedness hypothesis in financial markets. *Journal of Financial and Quantitative Analysis*, 30(01):23–42.

Caldana, R. and Fusai, G. (2013). A general closed-form spread option pricing formula. *Journal of Banking & Finance*, 37:4893ñ4906.

Carbonell, F., Jimenez, J., and Pedroso, L. (2008). Computing multiple integrals involving matrix exponentials. *Journal of Computational and Applied Mathematics*, 213(1):300–305.

Carmona, R. and Durrleman, V. (2003). Pricing and hedging spread options. *Siam Review*, 45(4):627–685.

Casassus, J. and Collin-Dufresne, P. (2005). Stochastic convenience yield implied from commodity futures and interest rates. *Journal of Finance*, 60(5):2283–2331.

Casassus, J., Liu, P., and Tang, K. (2013). Economic linkages, relative scarcity, and commodity futures returns. *Review of Financial Studies*, 26(5):2723–2749.

Comte, F. (1999). Discrete and continuous time cointegration. *Journal of Econometrics*, 88(2):207–226.

Cortazar, G., Milla, C., and Severino, F. (2008). A multicommodity model of futures prices: Using futures prices of one commodity to estimate the stochastic process of another. *Journal of Futures Markets*, 28(6):537–560.

Cortazar, G. and Naranjo, L. (2006). An N-factor gaussian model of oil futures prices. *Journal of Futures Markets*, 26(3):243ñ268.

Crowder, W. J. and Hamed, A. (1993). A cointegration test for oil futures market efficiency. *Journal of Futures Markets*, 13(8):933–941.

Dempster, M., Medova, E., and Tang, K. (2008). Long term spread option valuation and hedging. *Journal of Banking & Finance*, 32(12):2530–2540.

Duan, J.-C. and Pliska, S. R. (2004). Option valuation with co-integrated asset prices. *Journal of Economic Dynamics and Control*, 28(4):727–754.

Duffie, D. and Kan, R. (1996). A yield-factor model of interest rates. *Mathematical Finance*, 6:379–406.

Engle, R. F. and Granger, C. W. (1987). Co-integration and error correction: representation, estimation, and testing. *Econometrica*, pages 251–276.

Geman, H. and Ohana, S. (2009). Forward curves, scarcity and price volatility in oil and natural gas markets. *Energy Economics*, 31:576 – 585.

Gibson, R. and Schwartz, E. S. (1990). Stochastic convenience yield and the pricing of oil contingent claims. *The Journal of Finance*, 45(3):959–976.

Hambly, B., Howison, S., and Kluge, T. (2009). Modelling spikes and pricing swing options in electricity markets. *Quantitative Finance*, 9(8):937–949.

Johansen, S. (1991). Estimation and hypothesis testing of cointegration vectors in gaussian vector autoregressive models. *Econometrica*, pages 1551–1580.

Kirk, E. (1995). Correlation in the energy markets. *Managing energy price risk*, pages 71–78.

Kolos, S. P. and Ronn, E. I. (2008). Estimating the commodity market price of risk for energy prices. *Energy Economics*, 30:621–641.

Li, M., Zhou, J., and Deng, S.-J. (2010). Multi-asset spread option pricing and hedging. *Quantitative Finance*, 10(3):305ñ324.

Lo, A. and Wang, J. (1995). Implementing option pricing models when asset returns are predictable. *Journal of Finance*, 50:87ñ129.

Lucia, J. J. and Schwartz, E. S. (2002). Electricity prices and power derivatives: Evidence from the nordic power exchange. *Review of Derivatives Research*, 5:5–50.

Manoliu, M. and Tompaidis, S. (2002). Energy futures prices: term structure models with Kalman filter estimation. *Applied Mathematical Finance*, 9:21 – 43.

Margrabe, W. (1978). The value of an option to exchange one asset for another. *The Journal of Finance*, 33(1):177–186.

Meucci, A. (2009). Review of statistical arbitrage, cointegration, and multivariate Ornstein-Uhlenbeck. Working paper.

Nakajima, K. and Ohashi, K. (2012). A cointegrated commodity pricing model. *Journal of Futures Markets*, 32(11):995–1033.

Paschke, R. and Prokopczuk, M. (2009). Integrating multiple commodities in a model of stochastic price dynamics. *Journal of Energy Markets*, 2(3):47ñ82.

Paschke, R. and Prokopczuk, M. (2010). Commodity derivatives valuation with autoregressive and moving average components in the price dynamics. *Journal of Banking & Finance*, 34:2742 – 2752.

Phillips, P. (1991). Error correction and long-run equilibrium in continuous time. *Econometrica*, 59(4):967–980.

Schwartz, E. and Smith, J. E. (2000). Short-term variations and long-term dynamics in commodity prices. *Management Science*, 46(7):893–911.

Sørensen, C. (2002). Modeling seasonality in agricultural commodity futures. *Journal of Futures Markets*, 22(5):393ñ426.

Van der Werf, K. W. (2007). Covariance of the ornstein-uhlenbeck process. *Personal Communication*.

A Stationarity of $\Theta_z \boldsymbol{Z}$

The process $\boldsymbol{Z}$ is not stationary but there exist $n+h$ cointegration relationships between its components. We show in this appendix that the vector process $\Theta_z \boldsymbol{Z}$ is stationary. By construction, the first $n+h$ rows of the matrix Θ_z are non-zero row vectors and the $n+h$ non-zero eigenvalues of K have positive real part. $\Theta_z \boldsymbol{Z}$ has the following dynamics:

$$d(\Theta_z \boldsymbol{Z}(t)) = [\Theta_z \boldsymbol{\mu} - \Theta_z K \boldsymbol{Z}(t)]dt + \Theta_z \Sigma^{\frac{1}{2}} d\boldsymbol{W}(t).$$

$\Theta_z \boldsymbol{Z}(t)$ has a normal conditional distribution with two first moments:

$$\mathbb{E}[\Theta_z \boldsymbol{Z}(T) | \Theta_z \boldsymbol{Z}(t) = \Theta_z \boldsymbol{z}_t] = e^{-\Theta_z K_z (T-t)} \Theta_z \boldsymbol{z}_t + \left[\int_t^T e^{-\Theta_z K_z (T-u)} du \right] \Theta_z \boldsymbol{\mu}$$
$$Cov(\Theta_z \boldsymbol{Z}(T) | \Theta_z \boldsymbol{Z}(t)) = \int_t^T e^{-\Theta_z K_z (T-u)} \Theta_z \Sigma \Theta_z^T e^{-\Theta_z K_z^T (T-u)} du.$$

Let us focus on the non-zero components of $\Theta_z \boldsymbol{Z}$ ($n+h$ elements). The matrix $\Theta_z K_z$ is block diagonal with the only non-zero component in the left $(n+h) \times (n+h)$ corner. Noting that the notions of weak and strong stationarity are equivalent in a Gaussian framework, the non-zero components of $\Theta_z \boldsymbol{Z}$ are stationary if their first two limiting conditional moments are finite. This is the case if all the eigenvalues of the matrix in the left $(n+h) \times (n+h)$ corner of $\Theta_z K_z$ have positive real part. But the eigenvalues of the matrix in the left $(n+h) \times (n+h)$ corner are the non-zero eigenvalues of $\Theta_z K_z$ which are the non-zero $n+h$ eigenvalues of $K_z \Theta_z = K$. Therefore, $\Theta_z \boldsymbol{Z}$ is a stationary process.

B Characteristic function of $\boldsymbol{X}(t+\tau)$

Lemma 1. *The characteristic function of $\boldsymbol{X}(t+\tau)$ conditional on the information up to and including time t, i.e.,*

$$\varphi_{\boldsymbol{X}}(t, \tau, x, y; \omega) := \mathbb{E}_t \left[\exp \left\{ i\omega^\top \boldsymbol{X}(t+\tau) \right\} | \boldsymbol{X}(t) = x, \boldsymbol{Y}(t) = y \right], \tag{24}$$

is given by[21]

$$\begin{aligned}\varphi_{\boldsymbol{X}}(t,\tau,x,y;\omega) = \exp\Big\{ & i\omega^\top e^{-K_x\tau}x + i\omega^\top\psi(\tau)y + i\omega^\top\left[\phi(t+\tau) - e^{-K_x\tau}\phi(t)\right] \\ & + i\omega^\top\left[\int_0^\tau \psi(\tau-u)du\right]\boldsymbol{\mu}_y \\ & - \frac{1}{2}\begin{bmatrix}\omega^\top & \mathbf{0}_n^\top\end{bmatrix}\left[e^{-K\tau}\left(\int_0^\tau e^{Ku}\Sigma e^{K^\top u}du\right)e^{-K^\top\tau}\right]\begin{bmatrix}\omega \\ \mathbf{0}_n\end{bmatrix}\end{aligned} \tag{25}$$

Proof. By straightforward calculations. **q.e.d.**

C Equivalent convenience yield model

Similarly to Schwartz and Smith (2000), our model can also be written as a convenience yield model. Adopting the notation and structure of convenience yield models (e.g. Gibson and Schwartz (1990), Nakajima and Ohashi (2012)), we denote by $\boldsymbol{\delta}(t)$ the time-t vector of convenience yields and rewrite the dynamics of $\boldsymbol{X}(t)$ as follows:

$$d\boldsymbol{X}(t) = [\boldsymbol{\mu}_x - \boldsymbol{\delta}(t) - K_y\Theta\boldsymbol{X}(t)]dt + \Sigma_x^{\frac{1}{2}}d\boldsymbol{W}_x(t) \tag{26}$$

By matching the drift terms of (26) and (1), we obtain

$$\begin{aligned}\boldsymbol{\delta}(t) &= \boldsymbol{\mu}_x + (K_x - K_y\Theta)(\boldsymbol{X}(t) - \boldsymbol{Y}(t)) - K_y\Theta\boldsymbol{Y}(t) \\ &= \boldsymbol{\mu}_x + \begin{bmatrix}K_x - K_y\Theta, & -K_x\end{bmatrix}\boldsymbol{Z}(t)\end{aligned}$$

Note that the seasonality term $\phi(t)$ has been suppressed for simplicity.

It follows that $\boldsymbol{\delta}(t)$ is a linear combination of $\boldsymbol{X}(t) - \boldsymbol{Y}(t)$ and $\Theta\boldsymbol{Y}(t)$, which represent the cointegration relations and thus are stationary. Hence $\boldsymbol{\delta}(t)$ is stationary. Its dynamics is given by:

$$d\boldsymbol{\delta}(t) = [\boldsymbol{\mu}_\delta - K_\delta\boldsymbol{\delta}(t) - K_{\delta x}\Theta\boldsymbol{X}(t)]dt + \Sigma_{x\delta}^{\frac{1}{2}}d\boldsymbol{W}_x(t) + \Sigma_\delta^{\frac{1}{2}}d\boldsymbol{W}_y(t),$$

[21] The integrals appearing here can be computed explicitly using results in Carbonell et al. (2008).

where

$$\begin{aligned}\boldsymbol{\mu}_\delta &= -K_x\boldsymbol{\mu}_y + (K_xK_y\Theta - K_y\Theta K_x)K_x^{-1}\boldsymbol{\mu}_x\\ K_\delta &= K_x + (K_xK_y\Theta - K_y\Theta K_x)K_x^{-1}\\ K_{\delta x} &= (K_xK_y\Theta - K_y\Theta K_x)K_x^{-1}K_y\end{aligned}$$

For brevity we omit the explicit expressions for the variance-covariance matrices Σ_δ and $\Sigma_{x\delta}$, but, as in the case of $\boldsymbol{\mu}_\delta$, K_δ and $K_{\delta x}$, they depend on the parameters of the original model.

The following observations are worth mentioning: • As to be expected, the convenience yield model is also characterized by two factors, i.e. the vector of log-prices $\boldsymbol{X}(t)$ and the vector of convenience yields $\boldsymbol{\delta}(t)$. • While the convenience yields are stationary, the log-prices are not stationary, but are cointegrated. The cointegration matrix is given by Θ. • The vector of deviations from the cointegration relationships, i.e. $\Theta\boldsymbol{X}(t)$, is stationary and acts as an error-correction mechanism in the dynamics of log-prices, as well as convenience yields.

D Proof of Proposition 1

By the no-arbitrage assumption, the j-th futures price is given by

$$F_j(t,T) = \mathbb{E}_t^*\left[S_j(T)\right] = \mathbb{E}_t^*\left[\exp(X_j(T))|\boldsymbol{X}(t)=\boldsymbol{x},\boldsymbol{Y}(t)=\boldsymbol{y}\right] = \varphi^*_{\boldsymbol{X}}(t,\tau,x,y;-i\omega_j). \tag{27}$$

where $\varphi^*_{\boldsymbol{X}}(\cdot)$ is the characteristic function of $\boldsymbol{X}(T)$ under the risk-neutral measure, and where ω_j is a n-dimensional vector with the j-th component equal to 1 and the other components equal to 0. Noting that $\varphi^*_{\boldsymbol{X}}(\cdot)$ is similar to $\varphi_{\boldsymbol{X}}(\cdot)$ (see Lemma 1) but with $\boldsymbol{\mu}$ replaced by $\boldsymbol{\mu}^*$, we find

$$\begin{aligned}F_j(t,T) = \exp\Bigg\{&\omega_j^\top e^{-K_x\tau}\boldsymbol{X}(t) + \omega_j^\top\psi(\tau)\boldsymbol{Y}(t) + \omega_j^\top\left[\phi(t+\tau) - e^{-K_x\tau}\phi(t)\right]\\ &+\omega_j^\top\left[\int_0^\tau e^{-K_x(\tau-u)}du\right]\mu_x^* + \omega_j^\top\left[\int_0^\tau\psi(\tau-u)du\right]\mu_y^*\\ &+\frac{1}{2}\begin{bmatrix}\omega_j^\top & \mathbf{0}_n^\top\end{bmatrix}\left[e^{-K\tau}\left(\int_0^\tau e^{Ku}\Sigma e^{K^\top u}\right)e^{-K^\top\tau}\right]\begin{bmatrix}\omega_j\\ \mathbf{0}_n\end{bmatrix}\Bigg\}.\end{aligned}$$

The proposition can be verified now by a straightforward calculation.

E Estimation results

E.1 Parameter estimates

Below we report parameter estimates[22] for data set I, II and III under the convention $\widetilde{\boldsymbol{\lambda}} := \Sigma^{\frac{1}{2}}\boldsymbol{\lambda}$ and

$$\Sigma := \begin{bmatrix} diag(\sigma_x) & \mathbf{O}_n \\ \mathbf{O}_n & diag(\sigma_y) \end{bmatrix} \begin{bmatrix} \rho_x & \rho_{xy} \\ \rho_{xy}{}^{\top} & \rho_y \end{bmatrix} \begin{bmatrix} diag(\sigma_x) & \mathbf{O}_n \\ \mathbf{O}_n & diag(\sigma_y) \end{bmatrix}.$$

Dataset I:

$$\hat{\mu}_x = \begin{bmatrix} 0.00 \\ 0.00 \\ 0.00 \\ 0.00 \end{bmatrix}, \ \hat{\widetilde{\boldsymbol{\lambda}}}_x = \begin{bmatrix} 0.10^{***} \\ 0.08^{***} \\ 0.13^{***} \\ -0.10^{***} \end{bmatrix}, \ \hat{K}_x = \begin{bmatrix} 1.18^{***} & 0.01 & 0.02^{**} & 0.01 \\ 0.04 & 1.18^{***} & 0.01 & -0.01 \\ 0.00 & 0.03 & 1.05^{***} & 0.00 \\ 0.00 & 0.00 & 0.00 & 2.48^{***} \end{bmatrix},$$

$$\hat{\sigma}_x = \begin{bmatrix} 0.35^{***} \\ 0.34^{***} \\ 0.34^{***} \\ 0.60^{***} \end{bmatrix}, \ \hat{\rho}_x = \begin{bmatrix} 1.00 & 0.88^{***} & 0.86^{***} & 0.26^{***} \\ 0.88^{***} & 1.00 & 0.84^{***} & 0.34^{***} \\ 0.86^{***} & 0.84^{***} & 1.00 & 0.26^{***} \\ 0.26^{***} & 0.34^{***} & 0.26^{***} & 1.00 \end{bmatrix}, \ \hat{\mu}_y = \begin{bmatrix} 2.27^{***} \\ -1.88^{***} \\ -0.29 \\ -2.68^{**} \end{bmatrix}, \ \hat{\widetilde{\boldsymbol{\lambda}}}_y = \begin{bmatrix} 0.13^{***} \\ 0.12^{***} \\ 0.15^{***} \\ 0.05 \end{bmatrix},$$

$$\hat{K}_y = \begin{bmatrix} 0.64^{***} & -0.11^{***} & 0.00 & 0.00 \\ -0.60^{***} & 0.12 & 0.00 & 0.00 \\ -0.10 & -0.03 & 0.00 & 0.00 \\ -0.76^{**} & 0.00 & 0.00 & 0.00 \end{bmatrix}, \ \hat{\Theta} = \begin{bmatrix} 1.00 & -0.69^{***} & -0.28^{***} & 0.03^{***} \\ 0.36^{***} & 1.00 & -1.47^{***} & 0.00 \\ 0.00 & 0.00 & 0.00 & 0.00 \\ 0.00 & 0.00 & 0.00 & 0.00 \end{bmatrix},$$

$$\hat{\sigma}_y = \begin{bmatrix} 0.23^{***} \\ 0.19^{***} \\ 0.28^{***} \\ 0.24^{***} \end{bmatrix}, \ \hat{\rho}_y = \begin{bmatrix} 1.00 & 0.59^{***} & 0.56^{***} & 0.08 \\ 0.59^{***} & 1.00 & 0.51^{***} & 0.28^{***} \\ 0.56^{***} & 0.51^{***} & 1.00 & 0.16^{**} \\ 0.08 & 0.28^{***} & 0.16^{**} & 1.00 \end{bmatrix}, \ \hat{\rho}_{xy} = \begin{bmatrix} 0.48^{***} & 0.64^{***} & 0.50^{***} & 0.22^{***} \\ 0.58^{***} & 0.52^{***} & 0.53^{***} & 0.22^{***} \\ 0.53^{***} & 0.62^{***} & 0.29^{***} & 0.22^{***} \\ 0.20^{***} & 0.10^{**} & 0.14^{***} & 0.37^{***} \end{bmatrix},$$

$$\hat{\boldsymbol{\chi}}_1 = \begin{bmatrix} 0.00 \\ 0.03^{***} \\ -0.04^{***} \\ -0.01^{***} \end{bmatrix}, \ \hat{\boldsymbol{\chi}}_2 = \begin{bmatrix} 0.00 \\ -0.01^{***} \\ 0.02^{***} \\ 0.01^{***} \end{bmatrix}.$$

[22] Statistical significance at 1% indicated by ***, statistical significance at 5% indicated by ** and statistical significance at 10% indicated by *.

Dataset II:

$$\hat{\mu}_x = \begin{bmatrix} 0.00 \\ 0.00 \\ 0.00 \\ 0.00 \end{bmatrix},\ \hat{\tilde{\boldsymbol{\lambda}}}_x = \begin{bmatrix} -0.58^{***} \\ -0.05 \\ -0.08 \\ 0.07 \end{bmatrix},\ \hat{K}_x = \begin{bmatrix} 1.59^{***} & 0.39 & 0.20^{***} & -0.02 \\ -1.37^{**} & 2.55^{***} & -0.24 & 0.35^{*} \\ -4.51^{***} & 1.27^{***} & 4.95^{***} & -0.26 \\ -0.40 & 0.17 & -0.13 & 1.38^{***} \end{bmatrix},$$

$$\hat{\sigma}_x = \begin{bmatrix} 0.33^{***} \\ 0.51^{***} \\ 0.43^{***} \\ 0.23^{***} \end{bmatrix},\ \hat{\rho}_x = \begin{bmatrix} 1.00 & 0.25^{***} & 0.61^{***} & 0.24^{***} \\ 0.25^{***} & 1.00 & 0.25^{***} & 0.27^{***} \\ 0.61^{***} & 0.25^{***} & 1.00 & 0.27^{***} \\ 0.24^{***} & 0.27^{***} & 0.27^{***} & 1.00 \end{bmatrix},\ \hat{\mu}_y = \begin{bmatrix} 0.92^{***} \\ 0.87^{***} \\ 0.21 \\ 0.01 \end{bmatrix},\ \hat{\tilde{\boldsymbol{\lambda}}}_y = \begin{bmatrix} 0.27^{*} \\ 0.30 \\ 0.22^{*} \\ 0.11 \end{bmatrix},$$

$$\hat{K}_y = \begin{bmatrix} 0.78^{***} & 1.20^{***} & 0.25^{***} & 0.00 \\ -0.81^{***} & 0.00 & -0.05 & 0.00 \\ 0.24^{**} & 0.00 & 0.37^{***} & 0.00 \\ -0.17 & -0.19^{***} & 0.00 & 0.00 \end{bmatrix},\ \hat{\Theta} = \begin{bmatrix} 1.00 & -1.33^{***} & 0.00 & 0.00 \\ 0.00 & 1.00 & -0.63^{***} & 0.00 \\ 0.00 & 0.00 & 1.00 & -0.74^{***} \\ 0.00 & 0.00 & 0.00 & 0.00 \end{bmatrix},$$

$$\hat{\sigma}_y = \begin{bmatrix} 0.34^{***} \\ 0.42^{***} \\ 0.32^{***} \\ 0.29^{***} \end{bmatrix},\ \hat{\rho}_y = \begin{bmatrix} 1.00 & 0.72^{***} & 0.60^{***} & 0.38^{***} \\ 0.72^{***} & 1.00 & 0.42^{***} & 0.48^{***} \\ 0.60^{***} & 0.42^{***} & 1.00 & 0.24^{***} \\ 0.38^{***} & 0.48^{***} & 0.24^{***} & 1.00 \end{bmatrix},\ \hat{\rho}_{xy} = \begin{bmatrix} -0.01 & -0.10 & 0.02 & -0.14 \\ 0.39^{***} & 0.37^{***} & 0.45^{***} & 0.17^{**} \\ 0.32^{***} & 0.21 & 0.14^{***} & 0.08 \\ 0.25^{**} & 0.15^{*} & 0.22^{*} & 0.28^{***} \end{bmatrix},$$

$$\hat{\boldsymbol{\chi}}_1 = \begin{bmatrix} 0.14^{***} \\ 0.13^{***} \\ 0.20^{***} \\ -0.01^{***} \end{bmatrix},\ \hat{\boldsymbol{\chi}}_2 = \begin{bmatrix} -0.01 \\ 0.01 \\ 0.01 \\ 0.02^{***} \end{bmatrix}.$$

Dataset III:

$$\hat{\mu}_x = \begin{bmatrix} 0.00 \\ 0.00 \end{bmatrix},\ \hat{\tilde{\boldsymbol{\lambda}}}_x = \begin{bmatrix} -0.59^{***} \\ -0.36^{**} \end{bmatrix},\ \hat{K}_x = \begin{bmatrix} 0.94^{**} & 0.60^{**} \\ -4.15^{***} & 2.41^{***} \end{bmatrix},\ \hat{\sigma}_x = \begin{bmatrix} 0.30^{***} \\ 0.46^{***} \end{bmatrix},\ \hat{\rho}_x = \begin{bmatrix} 1.00 & 0.31^{***} \\ 0.31^{***} & 1.00 \end{bmatrix},$$

$$\hat{\mu}_y = \begin{bmatrix} -0.06 \\ 0.02 \end{bmatrix},\ \hat{\tilde{\boldsymbol{\lambda}}}_y = \begin{bmatrix} 0.17^{*} \\ 0.52^{***} \end{bmatrix},\ \hat{K}_y = \begin{bmatrix} 0.00 & 0.00 \\ -0.02 & 0.00 \end{bmatrix},\ \hat{\Theta} = \begin{bmatrix} 1.00 & -2.27^{***} \\ 0.00 & 0.00 \end{bmatrix},\ \hat{\sigma}_y = \begin{bmatrix} 0.18^{***} \\ 0.30^{***} \end{bmatrix},$$

$$\hat{\rho}_y = \begin{bmatrix} 1.00 & 0.61^{***} \\ 0.61^{***} & 1.00 \end{bmatrix},\ \hat{\rho}_{xy} = \begin{bmatrix} 0.25^{**} & -0.33^{***} \\ 0.44^{***} & 0.20 \end{bmatrix},\ \hat{\boldsymbol{\chi}}_1 = \begin{bmatrix} 0.12^{***} \\ 0.14^{***} \end{bmatrix},\ \hat{\boldsymbol{\chi}}_2 = \begin{bmatrix} -0.09^{***} \\ 0.01 \end{bmatrix}.$$

E.2 Estimation errors

Figure 17 compares the actual log-prices of futures on crude oil to the model-implied prices, for contracts with different times-to-maturities. The plain line overlaps almost perfectly with the crosses, which shows that the data is remarkably well fitted.

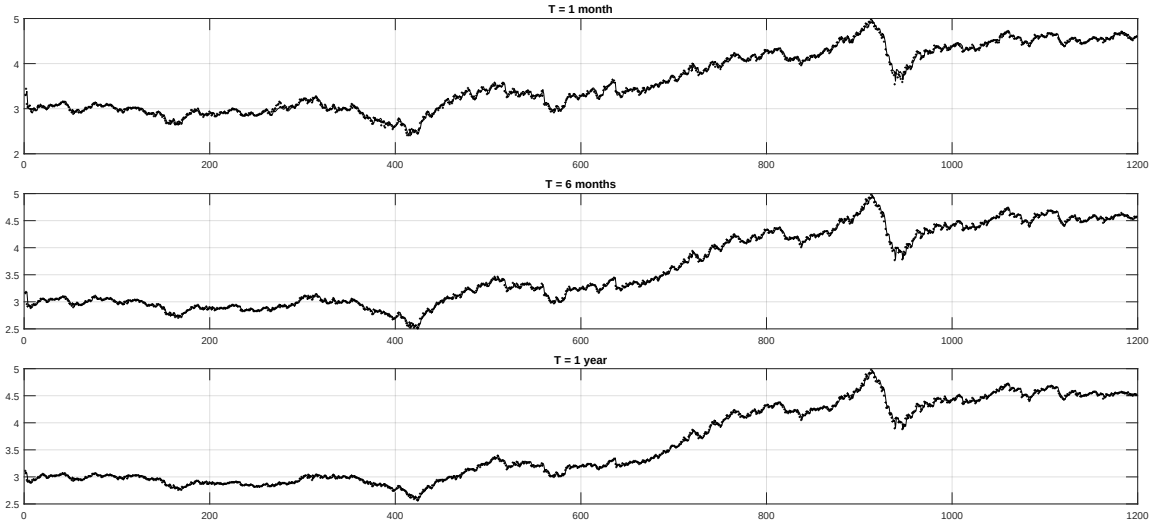

Figure 17. Comparison of the prices of log-futures on crude oil (crosses) with model-implied values (plain line) for contracts with different times-to-maturities ranging from one month (**top panel**) to one year (**bottom panel**).

Similar graphs are available upon request for other commodities.

E.3 Latent variables

Figures 18 through 27 show values of the filtered $\boldsymbol{X}(t)$ and $\boldsymbol{Y}(t)$. The bottom panels depict the seasonality patterns found in the data.

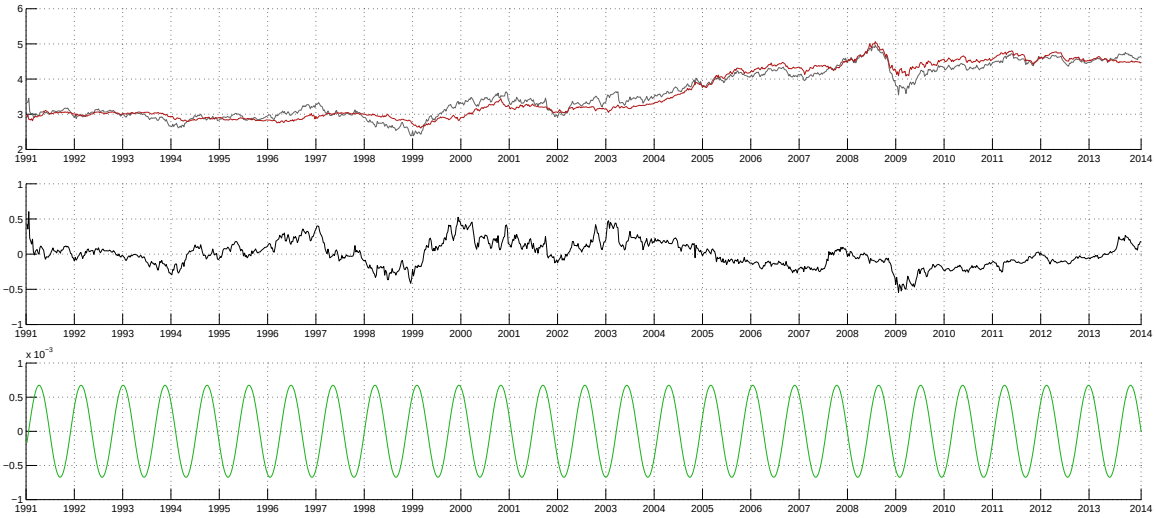

Figure 18. Top panel: Filtered time-series of spot log-prices $\boldsymbol{X}(t)$ (gray line) and long-run log-price level $\boldsymbol{Y}(t)$ (red line) for **crude oil**. **Middle panel:** Differences between short- and long-term log-price levels, i.e., $\boldsymbol{X}(t) - \phi(t) - \boldsymbol{Y}(t)$. **Bottom panel:** Estimated seasonality function for **crude oil**.

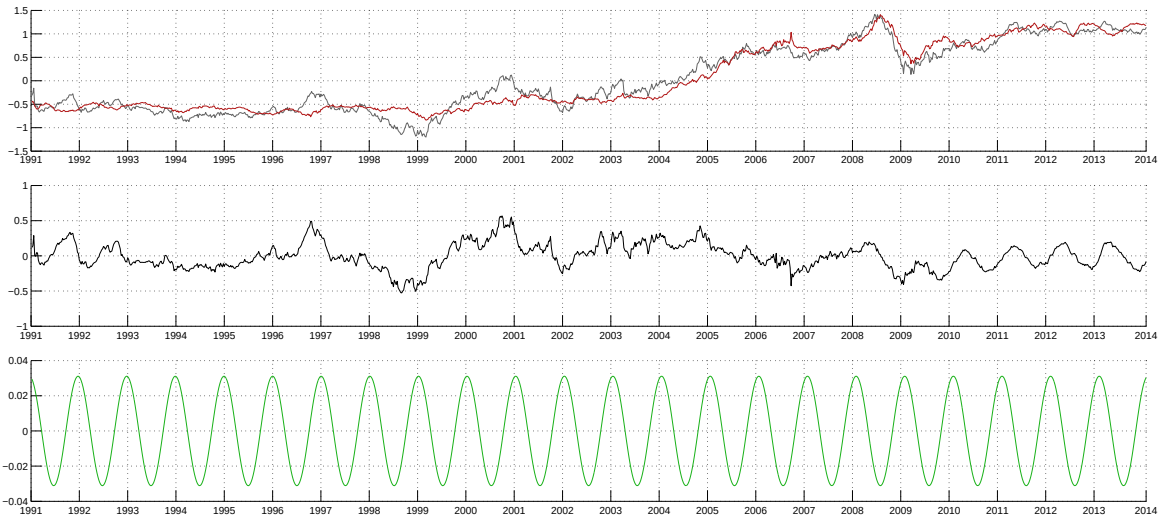

Figure 19. Top panel: Filtered time-series of spot log-prices $\boldsymbol{X}(t)$ (gray line) and long-run log-price level $\boldsymbol{Y}(t)$ (red line) for **heating oil**. **Middle panel:** Differences between short- and long-term log-price levels, i.e., $\boldsymbol{X}(t) - \phi(t) - \boldsymbol{Y}(t)$. **Bottom panel:** Estimated seasonality function for **heating oil**.

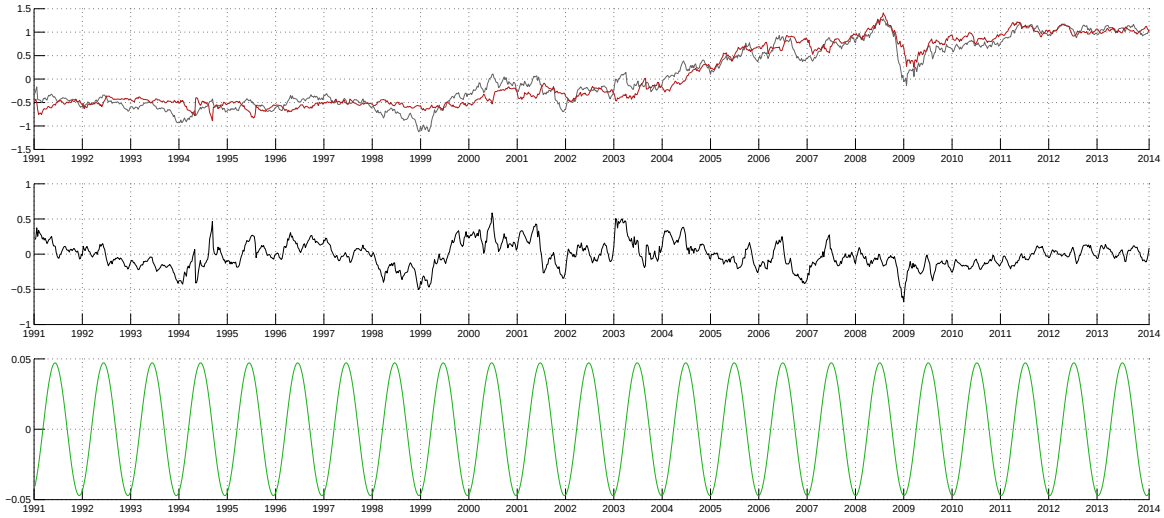

Figure 20. Top panel: Filtered time-series of spot log-prices $\boldsymbol{X}(t)$ (gray line) and long-run log-price level $\boldsymbol{Y}(t)$ (red line) for **gasoline**. **Middle panel:** Differences between short- and long-term log-price levels, i.e., $\boldsymbol{X}(t) - \phi(t) - \boldsymbol{Y}(t)$. **Bottom panel:** Estimated seasonality function for **gasoline**.

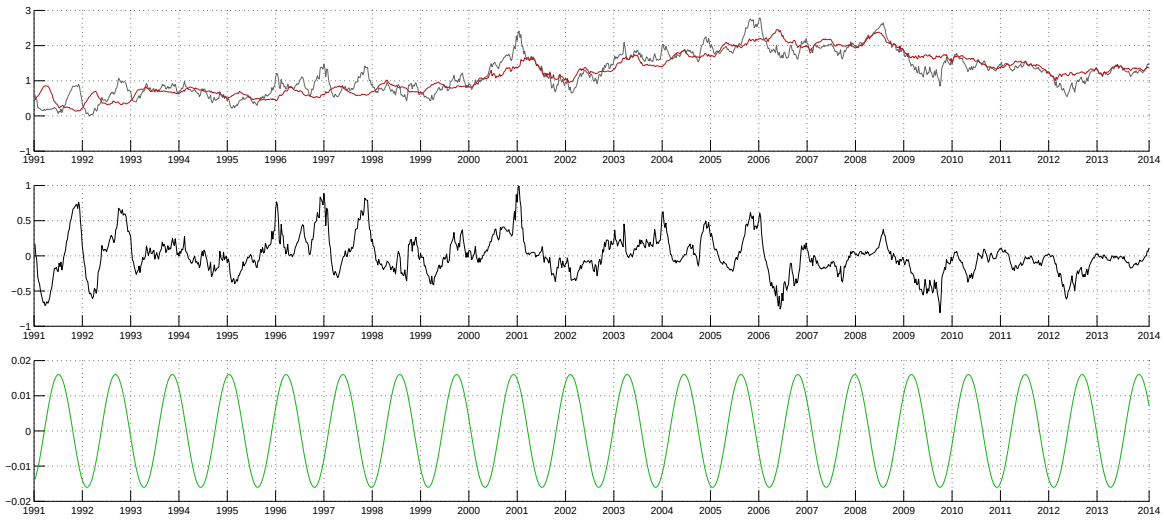

Figure 21. Top panel: Filtered time-series of spot log-prices $\boldsymbol{X}(t)$ (gray line) and long-run log-price level $\boldsymbol{Y}(t)$ (red line) for **natural gas**. **Middle panel:** Differences between short- and long-term log-price levels, i.e., $\boldsymbol{X}(t) - \phi(t) - \boldsymbol{Y}(t)$. **Bottom panel:** Estimated seasonality function for **natural gas**.

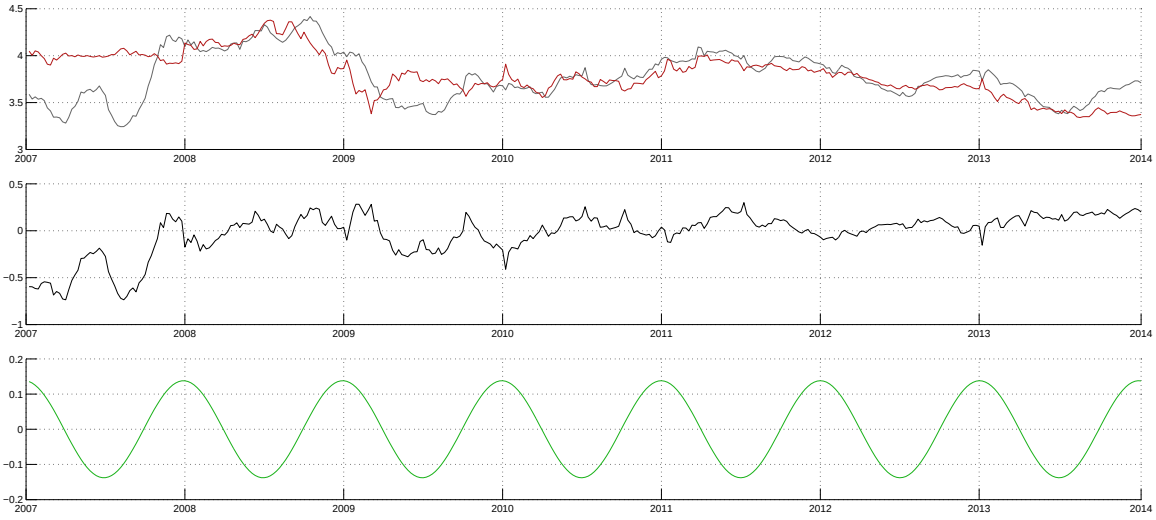

Figure 22. Top panel: Filtered time-series of spot log-prices $\boldsymbol{X}(t)$ (gray line) and long-run log-price level $\boldsymbol{Y}(t)$ (red line) for **power Germany**. **Middle panel:** Differences between short- and long-term log-price levels, i.e., $\boldsymbol{X}(t) - \phi(t) - \boldsymbol{Y}(t)$. **Bottom panel:** Estimated seasonality function for **power Germany**.

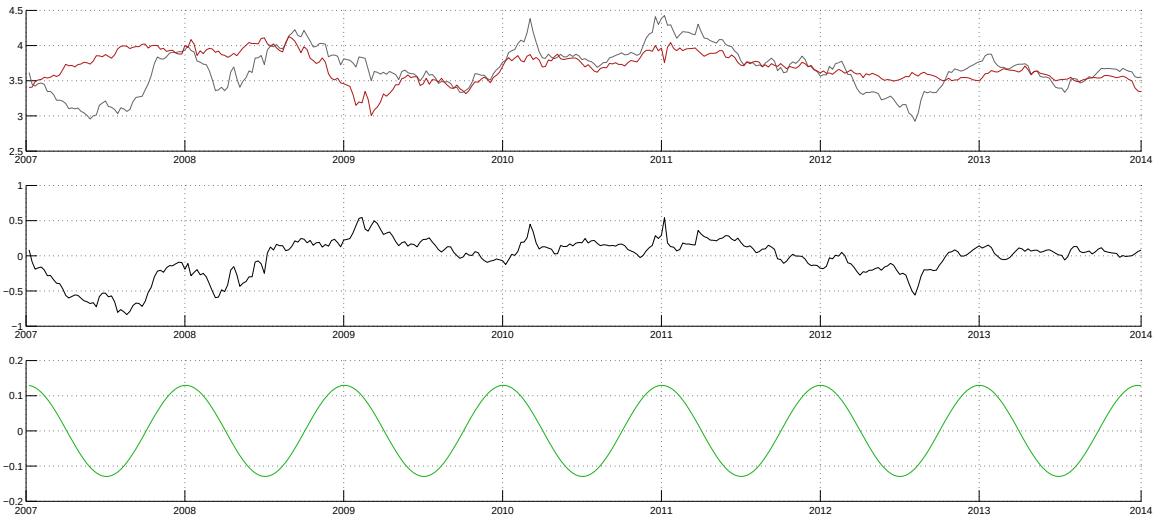

Figure 23. Top panel: Filtered time-series of spot log-prices $\boldsymbol{X}(t)$ (gray line) and long-run log-price level $\boldsymbol{Y}(t)$ (red line) for **power Nordics**. **Middle panel:** Differences between short- and long-term log-price levels, i.e., $\boldsymbol{X}(t) - \phi(t) - \boldsymbol{Y}(t)$. **Bottom panel:** Estimated seasonality function for **power Nordics**.

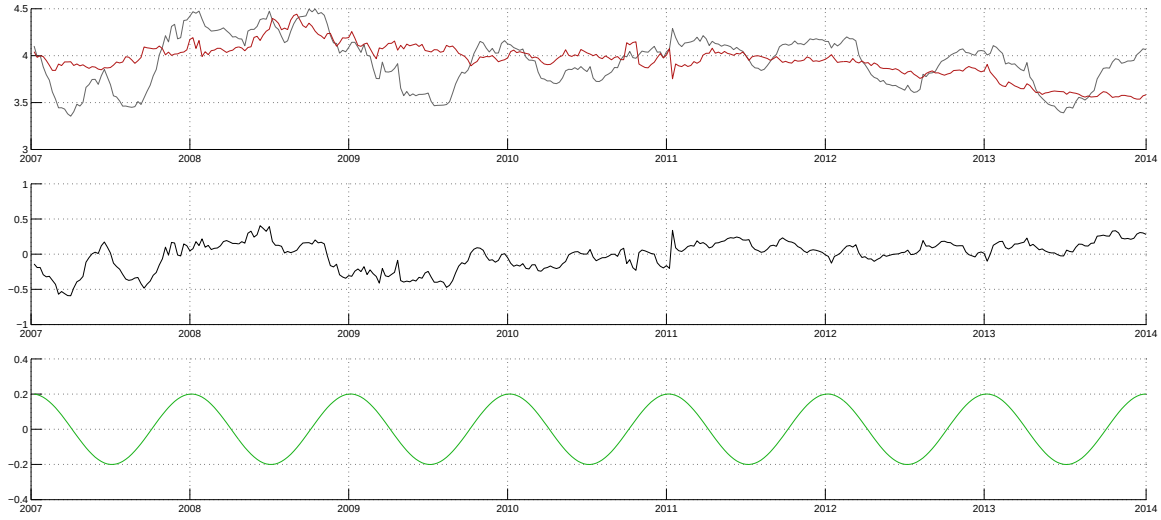

Figure 24. Top panel: Filtered time-series of spot log-prices $\boldsymbol{X}(t)$ (gray line) and long-run log-price level $\boldsymbol{Y}(t)$ (red line) for **power Swiss**. **Middle panel:** Differences between short- and long-term log-price levels, i.e., $\boldsymbol{X}(t) - \phi(t) - \boldsymbol{Y}(t)$. **Bottom panel:** Estimated seasonality function for **power Swiss**.

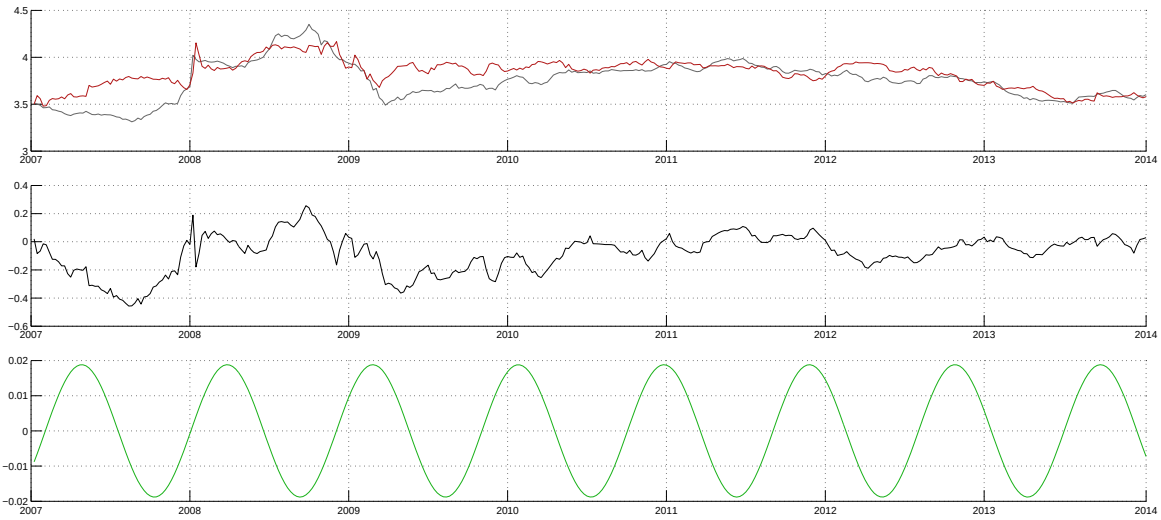

Figure 25. Top panel: Filtered time-series of spot log-prices $\boldsymbol{X}(t)$ (gray line) and long-run log-price level $\boldsymbol{Y}(t)$ (red line) for **power Poland**. **Middle panel:** Differences between short- and long-term log-price levels, i.e., $\boldsymbol{X}(t) - \phi(t) - \boldsymbol{Y}(t)$. **Bottom panel:** Estimated seasonality function for **power Poland**.

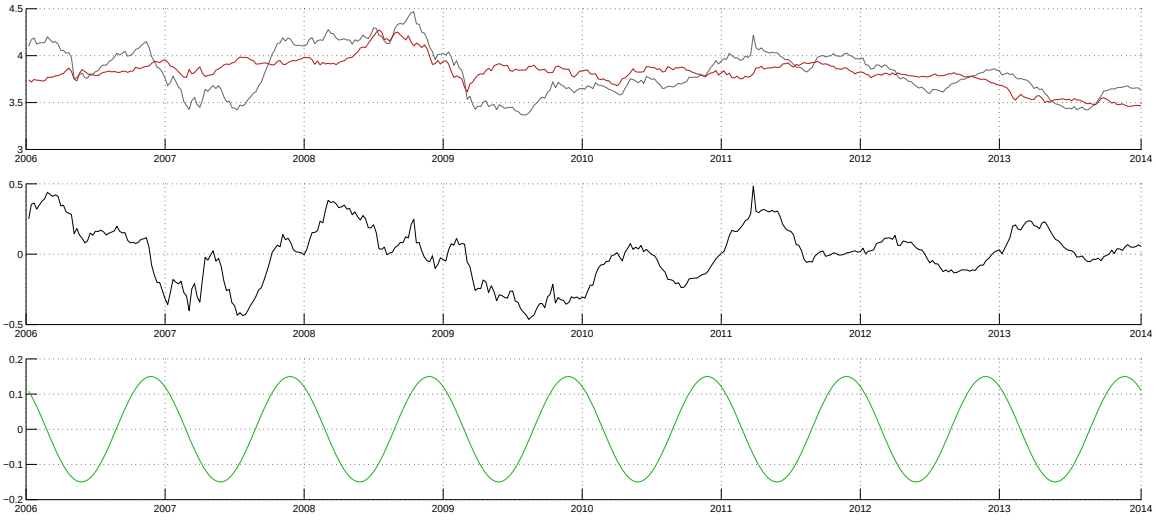

Figure 26. Top panel: Filtered time-series of spot log-prices $\boldsymbol{X}(t)$ (gray line) and long-run log-price level $\boldsymbol{Y}(t)$ (red line) for **electric power** (dataset III). **Middle panel:** Differences between short- and long-term log-price levels, i.e., $\boldsymbol{X}(t) - \phi(t) - \boldsymbol{Y}(t)$. **Bottom panel:** Estimated seasonality function for **electric power** (dataset III).

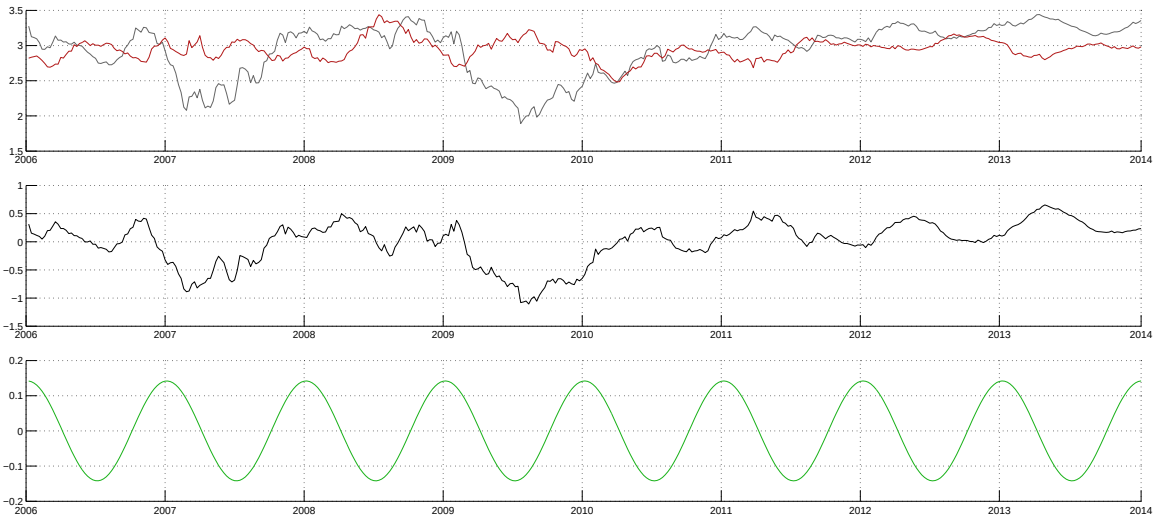

Figure 27. Top panel: Filtered time-series of spot log-prices $\boldsymbol{X}(t)$ (gray line) and long-run log-price level $\boldsymbol{Y}(t)$ (red line) for **natural gas** (dataset III). **Middle panel:** Differences between short- and long-term log-price levels, i.e., $\boldsymbol{X}(t) - \phi(t) - \boldsymbol{Y}(t)$. **Bottom panel:** Estimated seasonality function for **natural gas** (dataset III).